THE ASTROPHYSICAL JOURNAL, 173:557–592, 1972 May 1

GALACTIC SHOCKS IN AN INTERSTELLAR MEDIUM WITH TWO STABLE PHASES

FRANK H. SHU, VINCENZO MILIONE, AND WILLIAM GEBEL
State University of New York at Stony Brook

C. YUAN
City College of New York

D. W. GOLDSMITH*
University of California at Berkeley

AND

W. W. ROBERTS
University of Virginia

Received 1971 October 19; revised 1971 November 16

ABSTRACT

Quasi-steady flows of interstellar gas in a spiral gravitational field are followed for the purpose of investigating galactic shocks and the resultant processes of the formation of stars and interstellar clouds. We model the interstellar medium with two stable phases in which thermal balance is maintained through heating by low-energy cosmic rays. The problem, including transitions between the two phases, is given a general formulation but is solved in an approximation which ignores the difference in fluid velocities of the two phases. We also assume that the cosmic-ray flux is uniform in circles about the center of the Galaxy and that the relative abundances of the chemical elements are "normal."

For a spiral gravitational field with strength equal to 5 percent that of the axisymmetric field at 10 kpc from the galactic center, the density ratio at maximum and minimum compressions is 9:1 for the intercloud medium while it is 40:1 for the gas in a typical cloud. During the decompression phase of the flow, a small percentage of the mass of the clouds evaporates to become intercloud material, but this small amount is recovered in the shock. As a by-product of phase transitions, the properties of the clouds in the regions between spiral arms are such as to make their detection in 21-cm absorption very difficult.

In the absence of the cloud phase, we determine the thickness of the shock layer in the intercloud medium to be typically 50 pc. An interstellar cloud immersed as a test particle in the intercloud medium experiences a dynamic rather than a quasi-static compression as it passes through the shock layer. The critical mass for the gravitational collapse of a cloud is reduced by a large factor because of the compression in the shock.

I. INTRODUCTION

The problems discussed in this paper are motivated by the desire to understand the detailed mechanisms which trigger the formation of stars in normal spiral galaxies. Central to our discussion are two fundamental ideas: (i) spiral galactic shocks and (ii) the two-phase model of the interstellar medium. Within this context, we concentrate on the roles played by gravitational and thermal mechanisms. We avoid the vexing problem of the magnetic-field geometry by ignoring at the very outset the effects of the interstellar magnetic field. We do this not because we feel these effects to be unimportant, but because we wish to keep the present discussion as simple as possible.

a) Basic Concepts

On a small scale the main obstacle to star formation is that most of the interstellar clouds would not be even remotely bound by their self-gravitation *if the clouds were*

* Now at the State University of New York at Stony Brook.

isolated entities placed in a vacuum. This obstacle is largely removed by recent calculations which show that the nearly neutral component of the interstellar medium may consist of two gaseous phases in rough pressure equilibrium with one another (Zel'dovich and Pikel'ner 1969; Field, Goldsmith, and Habing 1969; Spitzer and Scott 1969; Silk and Werner 1969; Hjellming, Gordon, and Gordon 1969; Bottcher *et al.* 1970; Habing and Goldsmith 1971). These two phases are identified with the observed (dense) clouds at temperatures 20°–200° K and with an unobserved (rarefied) intercloud medium at temperatures $\sim 10^4$ ° K.

On a galactic scale, the problem of star formation presents the following puzzle. Why do the regions of star formation, as marked out by the giant H II regions in certain external galaxies, often appear to be lined up as "beads on a string" along two diametrically opposed spiral arms? A natural answer in the context of the density-wave theory of spiral structure is given by Roberts (1969; see also Fujimoto 1966, and Roberts and Yuan 1970). Roberts considers the nonlinear response of the interstellar gas to a spiral gravitational field of the type discussed by Lin, Yuan, and Shu (1969) and finds that shocks can develop. The large compression of interstellar gas behind the shocks is imagined to trigger the nearly simultaneous gravitational collapse of massive clouds along extensive fronts.

However, the previous discussions are incomplete, for the "effective pressure" in the gas is identified as resulting from the random motions of the interstellar clouds. For the large-scale dynamics, random motion can act as a "pressure" in that it provides support against the gravity field of the Galaxy (see § III); however, there are severe difficulties in visualizing how this "effective pressure" is transmitted on a small scale to trigger the gravitational collapse of clouds since cloud-cloud collisions provide compression essentially only in *one* direction (cf. Stone 1970).

The above difficulty is avoided if we adopt the two-phase model of the interstellar medium. In this picture, the galactic shocks are initiated by the intercloud medium; the "effective pressure" is simply the gas kinetic pressure of this medium; and the dissipation mechanisms by which the macroscopic energy of bulk motion is transformed to microscopic energy are the usual ones of viscosity, heat conduction, and radiation. In the shock layer the clouds are initially decelerated by the drag exerted by the intercloud medium and suffer further deceleration because cloud-cloud collisions serve as an effective means of (diffusive) momentum transport in regions where the gradient of the bulk velocity is large. Otherwise the clouds are viewed simply as embedded bodies which expand or contract to adjust to changes of the ambient pressure.

b) Procedure

The purpose of the present paper is to develop the theory for the description of the processes described above. In particular, we shall show that such a theory allows the inclusion of transitions between the two phases in a straightforward manner. This development is essential if we wish to understand the process by which interstellar clouds may be formed out of the intercloud medium.

Our investigation proceeds along the following line of attack. In § II we review briefly the basic features of the equilibrium model of an interstellar medium which contains two thermally stable phases heated by low-energy cosmic rays. Consideration of the various timescales shows that the equilibrium calculations are applicable locally to dynamical calculations everywhere except in a thin region—the shock layer—where the gas may briefly be carried far from an equilibrium state.

In § III and the Appendices we formulate the basic equations which govern the galactic flow. In particular, we explicitly give the form taken by the equations when phase transitions occur by the mechanism of thermal instability. In this section we treat galactic shocks as discontinuities in the flow for which appropriate jump conditions must be satisfied.

In § IV we solve these equations numerically in an approximation which ignores the difference in fluid velocities of the two phases. The most noteworthy result of this section is the conclusion that the probable state of the clouds in the regions *between spiral arms* is such as to make the clouds *extremely difficult to detect in 21-cm absorption.*

In § V we examine the internal structure of the shock layer in the intercloud medium. We do not consider here, as we do in § IV, the interactions with the cloud phase because the detailed knowledge of the relevant relaxation processes are not available at the present time. (For the calculations of § IV we need to know only that these processes exist and are rapid.) We do, however, make some preliminary comments on the behavior of the clouds in the shock layer when the clouds are treated as test particles.

In § VI we consider the gravitational stability of self-gravitating clouds which are bounded by external pressure. Clouds whose internal temperatures are maintained by external sources of heating have considerably greater compressibility than isothermal spheres, and we find the gravitational collapse of such clouds much easier to trigger by a sudden compression than has been previously estimated.

II. TWO-PHASE MODEL OF THE INTERSTELLAR MEDIUM

Because the H I gas of the interstellar medium is usually very rarefied, the excitation and ionization of atoms can occur either collisionally (via inelastic collisions with thermal particles or with energetic particles) or radiatively (via the absorption of starlight or of energetic photons) whereas de-excitation and recombination can occur only radiatively (via the spontaneous emission of a photon). The usual assumption of *local statistical equilibrium* for the population levels requires upward and downward transitions to be balanced in a fashion which is consistent with the local conditions of density and temperature.

The *microscopic* condition of equilibrium described above is not sufficient to guarantee a steady *macroscopic* state. In particular, if there is no external source to heat the gas, it will cool radiatively to very low temperatures in a time which is long compared with the microscopic relaxation times but short compared with the large-scale dynamical times encountered in the Galaxy (see § II*d*). We shall follow Field *et al.* (1969) and Spitzer and Scott (1969) in assuming the interstellar H I gas to be heated and kept partially ionized by a static flux of low-energy cosmic-ray protons. For simplicity, we shall assume the flux to be *uniform*, both inside and outside spiral arms. The latter assumption should be relaxed if the sources of heating and ionization should turn out, in fact, to be associated with extreme Population I objects.

As is evident from the number of papers in the literature dealing with this subject, great uncertainties are involved in estimating the flux of low-energy cosmic rays actually present in the general interstellar medium. Hence, it is safest perhaps to view the flux level to represent the mean input of heat by all sources of energetic particles and photons. Regarded in this spirit the flux level is a free parameter to be adjusted, within limits, to give a best fit for the observations. Fortunately, simple scaling laws exist to convert all of our computed results if the flux level is different from the adopted one.

a) Equilibrium Calculations

We assume that the input of heat is not so sporadic in time and in space that no meaning can be attached, even locally, to calculations of the equilibrium macroscopic state (cf. Bottcher *et al.* 1970). A steady state results per unit volume when the total heat gain Γ is balanced by the total radiative loss Λ, and when the total ionization rate I is balanced by the total recombination rate R (Eddington 1926).

Because the medium is assumed to be optically thin with respect to the cooling radiation, the equilibrium calculations can be represented, for given chemical composition, in terms of the variables (n, n_e, T, ζ). The notation used here is the usual one where n is the number density of all atoms, neutral and ionized, n_e is the number density

of free electrons; T is the kinetic temperature (assumed to be the same for all species of thermal particles); and ζ is the statistical rate of ionization of a single hydrogen atom by the *primary* flux of low-energy cosmic rays. Thus, the two steady-state conditions on heat and ionization balance can be written as the functional relations

$$\Gamma - \Lambda = \mathcal{H}(n, n_e, T, \zeta) = 0 , \tag{1a}$$

$$I - R = \mathcal{I}(n, n_e, T, \zeta) = 0 . \tag{1b}$$

In equations (1a) and (1b) we have suppressed explicit display of the dependences on the relative chemical abundances. The net heat gain per unit volume $\mathcal{H}$ is often written as $-\rho\mathcal{L}$, where $\mathcal{L}$ is the heat-loss function per unit mass, and ρ is the mass density given by

$$\rho = mn . \tag{2}$$

In the above m is the mean atomic mass *computed as if all the atoms were in their unionized states* and is equal to 2.16×10^{-24} g if the medium contains 10 percent helium by number.

The simultaneous solution of equations (1a) and (1b) allows us to express the pressure $P = (n + n_e)kT$ as a function P_{eq} of (ρ, ζ):

$$P = P_{eq}(\rho, \zeta) . \tag{3}$$

Plotted in figure 1 are the results of the steady-state calculation when ζ is taken to be 1.2×10^{-15} s^{-1}, and the relative chemical abundances are taken to be "normal."[1] The value $\zeta = 1.2 \times 10^{-15}$ s^{-1} is chosen to yield a density of free electrons in the intercloud medium which is in rough agreement with the density implied by pulsar dispersion measures (cf. Hjellming *et al.* 1969).

The high-density, low-temperature phase (clouds) can exist in pressure equilibrium with the low-density, high-temperature phase (intercloud gas) if $P = (n + n_e)kT$ falls between P_{max} and P_{min}. (Note from fig. 1 that $\log[P_{max}/k] = 3.18$ and $\log[P_{min}/k] = 2.48$.) The middle phase is one of unstable equilibrium. A bubble of gas placed artificially in this regime will tend to evolve toward one of the thermally stable phases on a time scale of a few million years (Goldsmith 1970). If the bubble is sufficiently small that pressure equilibrium with the ambient medium is easily established but not so small that thermal conductivity is important, the mode of thermal instability is isobaric (Field 1965).

b) Transitions between the Two Stable Phases

It is obvious that if one tries to increase the pressure above P_{max}, some of the intercloud material will condense to become cloud material, whereas if one tries to decrease the pressure below P_{min}, some of the cloud gas will evaporate to become intercloud gas. This process can only be loosely termed a thermal instability inasmuch as the medium is *driven* to a condition of runaway thermal imbalance by the disappearance of one of the stable equilibrium states. The phase transition in this case needs not go to completion because the transformation of some of the material into the other phase may relieve the pressure discrepancy for the remaining material.

[1] Apart from some minor exceptions, figure 1 is computed with the radiative loss functions used by Goldsmith, Habing, and Field (1969) and with the heating and ionization formulae given by Spitzer and Scott (1969). The differences are that we have followed Penston (1970) in using Peterson and Strom's (1969) rate coefficient for the collisional excitation of hydrogen (leading to the subsequent emission of Lα) and in including the cooling line λ5201 of N I. We have also adopted a number abundance of Fe relative to H of 3×10^{-5}. Each of these revisions increases the cooling efficiency of the intercloud medium relative to that used by Goldsmith *et al.* and largely accounts for the increase in a factor 3 in ζ which we find necessary to give essentially the same fit to observations.We have also included the process of collisional ionization of atomic hydrogen by electron impact. This process competes with cosmic-ray ionization at high temperatures and is thus not negligible in galactic shock layers.

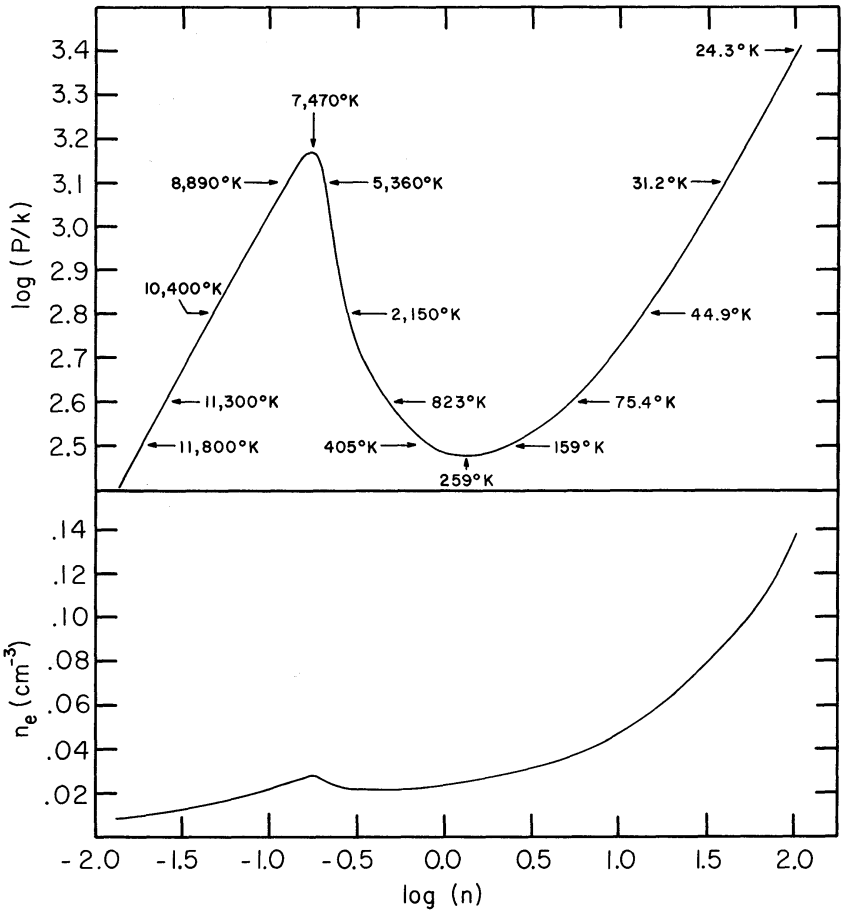


FIG. 1.—Equilibrium properties of interstellar gas heated by a uniform flux of low-energy cosmic-ray protons.

The above mechanism is one of *volume* radiation and is therefore a rapid transition if it occurs. Zel'dovich and Pikel'ner (1969) and Penston and Brown (1970) have also considered *surface* effects due to heat conduction at a cloud-intercloud interface. These considerations reveal a remarkable similarity between this gas and a van der Waals gas in the existence of an integral criterion similar to Maxwell's construction for the determination of the "vapor (or saturation) pressure" P_v. Phase transitions which arise at the interface when $P \neq P_v$ are, however, exceedingly slow if one adopts atomic transport coefficients; and we shall ignore this effect in a first treatment. Inside shock layers, however, the dynamical effects to be discussed in § V*d* may greatly enhance the effective transport coefficients and may lead to an ability for the preexisting clouds to act as "nucleation centers" for the transformation of intercloud material into cloud material.

c) Scaling Law for the Equilibrium Calculation

If we ignore plasma processes, then all the atomic processes which enter in the calculation of the quantities $\mathcal{H}$ and $\mathcal{G}$ are two-body processes. Indeed, an examination of the individual terms adopted in $\mathcal{H}$ and $\mathcal{G}$ shows that they are all proportional to n^2, nn_e, n_e^2, ζn, or ζn_e multiplied by some rate coefficient (and relative abundance) which depends only on T and possibly on the ratio n_e/n. Hence, as has been implicitly recognized by Field *et al.* (1969), the functions $\mathcal{H}$ and $\mathcal{G}$ have the equidimensional property

$$\mathcal{H}(\lambda n, \lambda n_e, T, \lambda\zeta) = \lambda^2\mathcal{H}(n, n_e, T, \zeta) , \tag{4a}$$

$$\mathcal{G}(\lambda n, \lambda n_e, T, \lambda\zeta) = \lambda^2 \mathcal{G}(n, n_e, T, \zeta) , \tag{4b}$$

where λ is any real positive number. Thus, from the solution for $\mathcal{H} = \mathcal{G} = 0$ given in figure 1 for $\zeta = 1.2 \times 10^{-15}\ \mathrm{s}^{-1}$, we may obtain solutions for any other value of ζ merely by scaling n and n_e and leaving T unchanged. In particular, the equilibrium pressure has the linear scaling property

$$P_{\mathrm{eq}}(\lambda\rho, \lambda\zeta) = \lambda P_{\mathrm{eq}}(\rho, \zeta) . \tag{5}$$

We shall make frequent use of these exact scaling relations throughout this paper. We also note from figure 1 that the two stable phases have P-ρ relations which are fairly well represented by the polytropic relation

$$P_{\mathrm{eq}}(\rho, \zeta) = K(\zeta)\rho^{(N+1)/N} , \tag{6}$$

where $K(\zeta) \propto \zeta^{-1/N}$ is a constant for given ζ. Table 1 shows that except for P very close to $P_{\max}$, the single value $N = -4.0$ suffices to give a very good fit for the intercloud phase and is used in § IV and in Appendix C to simplify calculations. In contrast, a constant value for N suffices for the cloud phase only if we restrict ourselves to density ranges of a factor $\sim$5 for P near $P_{\min}$ and of a factor $\sim$10 for P near $P_{\max}$. This proves sufficient to serve as a useful device for the calculation of the structure of self-gravitating clouds in § VI.

d) Applicability of the Equilibrium Calculations for Dynamics

The applicability of the equilibrium relations locally for dynamical calculations hinges on an examination of the various timescales. We list in table 2 some of the timescales relevant to our calculations.

In both the intercloud and cloud phases there exists a hierarchy of timescales: $t_{\mathrm{dyn}} \gg t_{\mathrm{recomb}} \gtrsim t_{\mathrm{cool}} \gg t_{\mathrm{atomic\ coll}}$. Furthermore, the relaxation time for a particular process in the intercloud medium is always much longer than the corresponding time in the cloud gas. The existence of such a hierarchy of timescales results in a great simplification for the development which is to follow.[2] Thus, except in a shock layer, the variations in pressure generally occur so slowly in the intercloud medium that the gas can easily adjust its local thermal properties to correspond to the equilibrium state. Moreover,

TABLE 1

POLYTROPIC INDICES BASED ON THE INSTANTANEOUS SLOPE OF THE EQUILIBRIUM PRESSURE-DENSITY RELATION

	INTERCLOUD PHASE		CLOUD PHASE	
P/k (cm^{-3} ° K)	n (cm^{-3})	$\{[(\partial \log P_{\mathrm{eq}})/(\partial \log \rho)]_\zeta - 1\}^{-1}$	n (cm^{-3})	$\{[(\partial \log P_{\mathrm{eq}})/(\partial \log \rho)]_\zeta - 1\}^{-1}$
300.....	0.0170	−3.78	1.20	−1.03
450.....	0.0292	−4.04	7.22	−1.75
600.....	0.0432	−4.14	12.8	−2.30
900.....	0.0740	−3.96	24.7	−2.96
1200.....	0.111	−3.22	37.7	−3.37
1500.....	0.164	−1.33	51.4	−3.66
5000.....	...	...	247.	−5.00

[2] Our discussion is physical, but it can be made rigorous by the introduction of the formal apparatus of "two-timing"—or, since we deal with steady flows, by the introduction of "multiple length scales."

TABLE 2

TYPICAL TIMESCALES IN THE TWO-PHASE COMPONENT MODEL OF THE INTERSTELLAR MEDIUM

Dynamic		Thermal	
Propagation of *forced* pressure disturbances of galactic scale in the intercloud medium (years)	$\sim 10^8$	Recombination and cooling (years)*:	
Sound travel time through clouds (years)	$\sim 10^7$	Intercloud	10^6–10^7
Cloud-cloud collisions (years)	$\sim 10^7$	Cloud	10^4–10^5
		Atomic collisions (years)†:	
		Intercloud	10^2–10^3
		Cloud	10^1–10^2

* Recombination times are generally somewhat longer than cooling times.

† The times required for the collisional excitation of neutral atoms and of ions by impacts with thermal electrons and with neutral hydrogen atoms are generally somewhat longer than the typical mean free-flight time for elastic collisions.

$(t_{\rm dyn})_{\rm intercloud} \gg (t_{\rm sound})_{\rm cloud} \sim t_{\rm cloud\text{-}cloud\ coll}$ so that the clouds can remain in rough pressure equilibrium with the surrounding intercloud medium. On the other hand, $t_{\rm cloud\text{-}cloud\ coll} \gg (t_{\rm cool})_{\rm cloud}$ so that the clouds can be heated only transiently by cloud-cloud collisions (cf. Stone 1970). When large gradients are present to throw the gases temporarily out of equilibrium, relaxation back to equilibrium is much slower for the macroscopic thermal state than for the microscopic degrees of freedom.

III. BASIC EQUATIONS

Two-phase flows have been studied extensively in connection, for example, with foams, aerosols, emulsions, and blood. A rigorous derivation of the fluid equations of such systems has recently been given by Drew (1971). Our treatment below and in Appendix A follows a more physical approach.

We begin with a description of the average properties of the clouds which is similar locally to a model of "standard clouds" (cf. Spitzer 1968). We assume that self-gravitation plays only a small role in binding most of the clouds. The density ρ_c inside the clouds will generally vary from cloud to cloud because of cloud-cloud collisions (among other things), but we assume that the mean value $\langle \rho_c \rangle$, obtained by averaging over the local distribution of clouds, is determined by the condition of pressure equilibrium:

$$P_{\rm eq}(\langle \rho_c \rangle, \zeta) = P \,, \tag{7}$$

where P is the pressure of the ambient intercloud medium.

Let F_c denote the local fraction of volume occupied by the clouds; n_s, the number of clouds per unit volume; and $\langle M \rangle$, the average mass of the clouds. Clearly, there exists the relation

$$F_c \langle \rho_c \rangle = n_s \langle M \rangle \,. \tag{8}$$

For convenience, we further adopt a model of spherical clouds and define the average cloud radius $\langle R_c \rangle$ through[3]

$$\frac{4\pi}{3} \langle \rho_c \rangle \langle R_c \rangle^3 = \langle M \rangle \,, \quad \text{or equivalently,} \quad F_c = n_s \frac{4\pi}{3} \langle R_c \rangle^3 \,. \tag{9}$$

We adopt a fluid description for the motions of the clouds in the Galaxy. We allow the possibility for transitions between the cloud and intercloud phases, but we assume

[3] The definition (9) can be justified only on the basis of convenience since there is no cogent reason why most of the clouds need be spherical if they are held together only by external pressure.

the net sink of gas due to star formation to be negligibly small. Let $\langle \dot{M} \rangle$ and $\langle \dot{M} \boldsymbol{v}' \rangle$ be the average time rates of change of the mass and momentum of a cloud due to phase transitions; $\boldsymbol{u}_c$, the mean velocity of the clouds; $\boldsymbol{u}$, the fluid velocity of the intercloud gas; c, the clouds' rms velocity dispersion in any direction; and $\mathcal{U}$, the gravitational potential of the Galaxy. The equations for mass and momentum transfer of the clouds are then given in an inertial frame of reference by

$$\frac{\partial}{\partial t}(F_c\langle \rho_c \rangle) + \nabla \cdot (F_c \langle \rho_c \rangle \boldsymbol{u}_c) = n_s \langle \dot{M} \rangle , \tag{10a}$$

$$\frac{\partial}{\partial t}(F_c\langle \rho_c \rangle \boldsymbol{u}_c) + \nabla \cdot (F_c \langle \rho_c \rangle \boldsymbol{u}_c \boldsymbol{u}_c) = -F_c \langle \rho_c \rangle \nabla \mathcal{U} - \nabla (F_c \langle \rho_c \rangle c^2) - F_c \nabla P - D(\boldsymbol{u}_c - \boldsymbol{u}) + n_s \langle \dot{M} \boldsymbol{v}' \rangle . \tag{10b}$$

For a derivation of these equations from a kinetic description of the clouds, see Appendix A.

The coefficient D associated with the drag force between the two phases is given by

$$D = \tfrac{3}{8} C F_c (1 - F_c) \rho \frac{|\boldsymbol{u}_c - \boldsymbol{u}|}{\langle R_c \rangle} , \tag{11}$$

where the dimensionless drag coefficient C can be estimated to be about unity for flows at Mach numbers above 0.5 (cf. Chernyi 1961). In what follows we regard ζ, $\langle M \rangle$, c, and $\mathcal{U}$ to be known. (In the usual model of standard clouds, $\langle M \rangle = 400\ M_\odot$, $c = 8$ km s^{-1}.) We further assume that the change of momentum due to phase transitions involves mass moving at velocity $\boldsymbol{u}_c$ if cloud material is being transformed into intercloud material and at velocity $\boldsymbol{u}$ if intercloud material is being transformed into cloud material (see Appendix C). Thus,

$$\begin{aligned} n_s \langle \dot{M} \boldsymbol{v}' \rangle &= n_s \langle \dot{M} \rangle \boldsymbol{u}_c \quad \text{if} \quad n_s \langle \dot{M} \rangle < 0 , \\ &= n_s \langle \dot{M} \rangle \boldsymbol{u} \quad \text{if} \quad n_s \langle \dot{M} \rangle > 0 . \end{aligned} \tag{12}$$

The fluid equations for the intercloud medium are easily written down as

$$\frac{\partial}{\partial t}[(1 - F_c)\rho] + \nabla \cdot [(1 - F_c)\rho \boldsymbol{u}] = -n_s \langle \dot{M} \rangle , \tag{13a}$$

$$\frac{\partial}{\partial t}[(1 - F_c)\rho \boldsymbol{u}] + \nabla \cdot [(1 - F_c)\rho \boldsymbol{u}\boldsymbol{u}] = -(1 - F_c)\rho \nabla \mathcal{U} - (1 - F_c)\nabla P - D(\boldsymbol{u} - \boldsymbol{u}_c) - n_s \langle M \boldsymbol{v}' \rangle . \tag{13b}$$

In equation (13b) we have ignored the turbulent contribution to the momentum transfer in the intercloud medium. Furthermore, under the conditions of local thermal equilibrium, the energy equation, the equation of charge balance, and the equation of state for the intercloud gas may be replaced by the P-ρ relation

$$P = P_{\text{eq}}(\rho, \zeta) . \tag{14}$$

Equations (7), (10), (13), and (14) cannot be used in galactic shock layers where departures from local thermal equilibrium exist in conjunction with steep gradients of density, velocity, pressure, and temperature. On a galactic scale, we treat such layers as discontinuities in the flow which may be incorporated by placing appropriate jump conditions on the fluid equations (see § III*c*).

For the present we note that equations (7)–(14) do not form a closed set even if ζ, $\langle M \rangle$, c, and $\mathcal{U}$ are assumed to be known. To close the set of equations, we need to specify $n_s \langle \dot{M} \rangle$ in terms of the existing variables. Fortunately, as we shall see in the next sub-

section, we are able to determine $n_s\langle\dot{M}\rangle$ for the large-scale problem without being required to make a detailed rate calculation of the physical processes which are involved. The reason is that the phase transitions of interest are those discussed in § II*b*, and the occurrence of these transitions is dominated by the conditions of the large-scale flow, rather than the reverse.

a) Form of the Equations with and without Phase Transitions

In a smoothly varying galactic flow involving the two stable phases described in § II, appreciable phase transition will occur only if the pressure in the intercloud medium tries to exceed $P_{\max}$ or to fall below $P_{\min}$. On the other hand, the thermal instability which sets in at this point occurs so quickly and is so efficient that it virtually supplies the excess or deficit in pressure required by the flow instantaneously by adjusting the ratio of cloud/intercloud gas. Hence, as long as the instability is operative and as long as the flow remains smooth (i.e., does not develop shocks), the pressure is regulated to have the value

$$P = P_m(\zeta) , \tag{15}$$

where P_m is $P_{\max}$ or $P_{\min}$ as the case may be. Conversely, as soon as the flow can reverse the trend of the change of pressure which brought on the phase transition, the instability ceases and phase transition is terminated. Of course, if the driving of the thermal instability by the flow is of sufficient duration, phase transition may go to completion and the pressure restriction (15) no longer applies.

In summary, either $n_s\langle\dot{M}\rangle = 0$ or $n_s\langle\dot{M}\rangle \neq 0$. If $n_s\langle\dot{M}\rangle = 0$, the set of equations (7)–(14) is closed. If $n_s\langle\dot{M}\rangle \neq 0$, the additional relation (15) closes the set.

b) Reduction to Two Dimensions

Because the interstellar gas is distributed in a very thin layer in the Galaxy, it is obviously advantageous to integrate the dynamical equations in the vertical direction. Such an integration removes terms involving differentiation with respect to z which would otherwise formally constitute the largest terms in the equations. It would be wrong, however, to infer that the three-dimensional nature of the distribution is intrinsically unimportant. For a given amount of matter, the mixture ratio of the two stable phases at a given pressure depends on the volume to which the system is confined. The most important direction of confinement is the vertical direction—a fact explicitly recognized by Field *et al.* (1969). In the present paper we are primarily interested in the physical consequences of periodic compressions and decompressions in the directions parallel to the galactic plane due to the presence of a spiral gravitational field. Since our qualitative conclusions will not depend sensitively on the exact details of the vertical distribution, we adopt in Appendix B the simplest self-consistent model for the vertical structure.

In the model considered in Appendix B, the z-dependence of $F_c\langle\rho_c\rangle$ is given by the Gaussian exp $(-\pi z^2/h_c^2)$ whereas the z-dependences of ρ, $\langle\rho_c\rangle$, and P are given by the Gaussian exp $(-\pi z^2/h^2)$. The quantities h_c and h are the "effective thicknesses" of the distributions of the two phases; their squares are given in terms of various quantities evaluated in the central plane of the Galaxy $z = 0$:

$$h_c^2 = \left(\frac{2\pi c^2}{\partial^2\mathcal{U}/\partial z^2}\right)_{z=0}, \qquad h^2 = \left(\frac{2\pi P/\rho}{\partial^2\mathcal{U}/\partial z^2}\right)_{z=0}. \tag{16}$$

To be definite, we also assume that the gravitational potential $\mathcal{U}$ corresponds to that of the density-wave theory of spiral structure and is (nearly) steady in a frame which rotates with angular velocity Ω_p about the z-axis. We look for flows which are

steady in this frame.[4] If the integrations of the fluid equations over z are now carried out asymptotically for small h_c and h (see Appendix B), the *two-dimensional* fluid equations now read

$$\nabla\cdot(F_c\langle\rho_c\rangle h_c\mathbf{u}_c) = n_s h_c\langle\dot{M}\rangle\,, \tag{17a}$$

$$\nabla\cdot[(1-F_c)\rho h\mathbf{u}] = -n_s h_c\langle\dot{M}\rangle\,, \tag{17b}$$

$$\nabla\cdot(F_c\langle\rho_c\rangle h_c\mathbf{u}_c\mathbf{u}_c) = F_c\langle\rho_c\rangle h_c[-\nabla(\mathcal{U}-\tfrac{1}{2}\Omega_p^2\varpi^2)-2\boldsymbol{\Omega}_p\times\mathbf{u}] - \nabla(F_c\langle\rho_c\rangle h_c c^2) - F_c\nabla(Ph) - Dh(\mathbf{u}_c-\mathbf{u}) + n_s h_c\langle\dot{M}\mathbf{v}'\rangle\,, \tag{17c}$$

$$\nabla\cdot[(1-F_c)\rho h\mathbf{u}\mathbf{u}] = (1-F_c)\rho h[-\nabla(\mathcal{U}-\tfrac{1}{2}\Omega_p^2\varpi^2)-2\boldsymbol{\Omega}_p\times\mathbf{u}] - (1-F_c)\nabla(Ph) - Dh(\mathbf{u}-\mathbf{u}_c) - n_s h_c\langle\dot{M}\mathbf{v}'\rangle\,. \tag{17d}$$

All the symbols, except for h_c and h, in equations (17) and in the equations which follow refer to their values in the plane $z = 0$.

c) *Jump Conditions*

On a galactic scale shocks are treated as discontinuities in the flow, and the equations (17) are to be supplemented by jump conditions if shocks are to be included. Actually, of course, the thickness of a shock layer in the direction s normal to the shock front is characterized by one (or more) small length scale, say s_0 (see § V). Inside the shock layer the flow variables may be divided into two classes: those which are $O(1)$ and those which are $O(s_0^{-1})$. Terms of the second type are $n_s\langle\dot{M}\rangle$, the normal component of the drag $Dh(u_{c\perp}-u_\perp)$, and the normal derivatives of F_c, $\langle\rho_c\rangle$, ρ, P, $u_{c\perp}$, and $u_\perp$, where $u_{c\perp}$ and $u_\perp$ are the components of velocity perpendicular to the shock front. The normal derivatives of the components $u_{c\parallel}$ and $u_\parallel$ parallel to the shock front are smoothly varying and are typical of terms of the first type.

Thus, the jump conditions to be imposed on equations (17) are obtained by integrating the equations across the shock layer in the direction normal to the front and taking the limit $s_0\to 0$. After a little manipulation, we obtain

$$[F_c\langle\rho_c\rangle h_c u_{c\perp}]_1^2 = \Delta_m\,, \tag{18a}$$

$$[(1-F_c)\rho h u_\perp]_1^2 = -\Delta_m\,, \tag{18b}$$

$$[F_c\langle\rho_c\rangle h_c(u_{c\perp}^2+c^2)+F_c Ph]_1^2 = \Delta_p\,, \tag{18c}$$

$$[(1-F_c)h(\rho u_\perp^2+P)]_1^2 = -\Delta_p\,. \tag{18d}$$

In the above the indices 1 and 2 denote positions in the Galaxy immediately upstream and immediately downstream from the shock. The quantities

$$\Delta_m = \int_1^2 n_s h_c\langle\dot{M}\rangle ds\,, \tag{19a}$$

$$\Delta_p = \int_1^2\left[n_s h_c\langle\dot{M}v_\perp'\rangle - Dh(u_{c\perp}-u_\perp)+Ph\frac{\partial F_c}{\partial s}\right]ds\,, \tag{19b}$$

[4] Field (private communication) has pointed out that galactic shocks are, strictly speaking, incompatible with Bernoulli's theorem for steady barotropic flow. If we adopt, however, the view of Lin and Shu (1964, 1966, 1972) and consider spiral structure to be intrinsically a quasi-stationary phenomenon, we avoid such difficulties by requiring only that the flow be quasi-steady. Our formal solutions are, then, properly regarded as being valid for times shorter than the characteristic time of evolution of the underlying spiral structure.

represent, respectively, the interchange of mass (between the two phases) due to phase transitions; and the interchange of momentum due to phase transitions, due to drag, and due to a difference in compressibility of the two phases.

By definition, the shock layer is that region of space where the system relaxes to local equilibrium conditions after it has been thrown out of a different equilibrium upstream. Thus, we require that the relation between P and ρ (and $\langle \rho_c \rangle$) immediately before and immediately after the shock correspond to the equilibrium conditions

$$P = P_{\rm eq}(\rho, \zeta) = P_{\rm eq}(\langle \rho_c \rangle, \zeta) \quad \text{at} \quad \text{"1" and "2."} \tag{20}$$

When phase transition is present inside the shock layer, there is an additional requirement on the flow. Because a shock is always compressional, i.e., $P(2) > P(1)$, any phase transition involved will be in the sense that intercloud material is being converted into cloud material. *This phase transition is forced on the flow because the compression in the absence of phase transition would imply downstream pressures $P(2)$ greater than $P_{\rm max}$.* On the other hand, it is unlikely that phase transition would continue once the pressure drops significantly below $P_{\rm max}$. *Therefore, in a shock which involves phase transition, we require that the downstream pressure $P(2)$ be close to (but perhaps slightly less than) $P_{\rm max}$* (see Appendix C):

$$P(2) \simeq P_{\rm max} . \tag{21}$$

Given the upstream values and the knowledge that phase transition is absent in the shock layer ($\Delta_m = 0$), equations (18) and (20) underspecify the downstream values of F_c, $\langle \rho_c \rangle$, ρ, $u_{c\perp}$, $u_\perp$, and P because we do not generally know the value of Δ_p. Similarly, equations (18), (20), and (21) underspecify the downstream conditions also by one degree of freedom. We shall now see that the one degree of freedom is removed by the steady-state requirements of the large-scale flow.

d) Self-Regulation of the Average Pressure in a Steady State

If we assume that the flow is (nearly) steady and that the streamlines are (nearly) closed, the values of Δ_m and Δ_p are specified by the requirement that there can be no net transfer of either mass or momentum between the two phases in a complete circuit around the Galaxy. Thus, the relaxation process inside the shock from "1" to "2" is required to produce amounts of mass transfer and momentum transfer which are exactly opposite those produced in the rest of the flow going the "long way" around the Galaxy from "2" to "1".

For a given spiral gravitational field, the specification of both Δ_m and Δ_p is not generally consistent with arbitrary "average" conditions in the Galaxy since there is only *one* degree of freedom associated with the jump conditions of § III*c*. For example, if the average pressure along a closed streamline is too high, we would expect to find that there is more conversion of intercloud material into cloud material in the shock than there is transformation of cloud material into intercloud material in the rest of the flow. This would, of course, result in a lowering of the average pressure because the material occupies less volume in the cloud form than in the intercloud form. Conversely, if the average pressure is too low, the net transformation of cloud material into intercloud material would tend to raise the average pressure. The requirement that there be no net momentum transfer between the two phases is, of course, guaranteed once we have succeeded in finding closed streamlines (cf. § IV*c*).

We believe the self-regulation described above to be the basic mechanism by which the conditions in the interstellar medium approach those of a steady state. Thus, in § IV, when rapid phase transitions via thermal instability are present, the average pressure is found to be uniquely determined by the condition that there be no net transformation of material along a (nearly) closed streamline. We remark that if the Galaxy were perfectly static on a large scale (no spiral structure), the average pressure

would *eventually* tend to the vapor pressure P_v because the only phase transition which can take place in this case is that discussed by Zel'dovich and Pikel'ner (1969) and by Penston and Brown (1970). This remark must also apply approximately if the spiral structure were very weak.

e) *Approximate Method of Solution*

So far the development has been rather formal. We now adopt an approximate method which greatly simplifies the solution of the dynamical equations. We assume that the velocities of the cloud and intercloud medium do not differ very much from one another except, possibly, immediately after a galactic shock in the intercloud medium (see § V*d*). This situation arises because the gravitational and inertial accelerations are the same for the clouds and the intercloud gas, and any differences between $\boldsymbol{u}_c$ and $\boldsymbol{u}$ which arise because of differences in the accelerations due to the "pressure" forces will be kept below a few kilometers per second by the drag.[5]

As an initial approximation, therefore, we set $\boldsymbol{u}_c = \boldsymbol{u}$. The adoption of this approximation means we must forgo the use of one of the momentum equations (17c) or (17d). (We must also forgo the possibility of following in detail the exchange of momentum between the two phases.) The bulk of the gaseous mass and momentum is contained in the cloud phase, but the bulk of the thermal energy is contained in the intercloud phase. In this paper we are primarily interested in following the variation of the gas kinetic pressure; therefore, we choose to use the momentum equation (17d) of the intercloud gas.

During compression of the entire medium, the intercloud gas tends to cool slightly (i.e., h tends to decrease slightly) to maintain thermal equilibrium, whereas $(1 - F_c)$ tends to increase somewhat because the clouds compress more easily than the intercloud gas. As another approximation, therefore, we ignore the variation of $(1 - F_c)h$ and h in equations (17b), (17d), (18b), and (18d). [It is found *a posteriori* that $(1 - F_c)h$ and h do not vary by more than 10 percent along a streamline with average radius 10 kpc from the galactic center.]

In the above approximations the equations of continuity (17a) and (17b) can be divided by $(1 - F_c)h$. *In the absence of phase transitions*, these equations now read

$$\nabla\cdot(f_c\langle\rho_c\rangle\boldsymbol{u}) = 0\,, \tag{22a}$$

$$\nabla\cdot(\rho\boldsymbol{u}) = 0\,, \tag{22b}$$

where we have written

$$f_c = \frac{F_c h_c}{(1 - F_c)h} \tag{23}$$

as the fraction of the surface area (projected volume) occupied by the clouds relative to that occupied by the intercloud medium. *In the presence of phase transitions*, we add the separate equations of continuity to eliminate $n_s h_c\langle\dot{M}\rangle$ and write

$$\nabla\cdot[(f_c\langle\rho_c\rangle + \rho)\boldsymbol{u}] = 0\,. \tag{24}$$

In a similar way, if we use the equation of continuity for the intercloud medium to eliminate the term $n_s h_c\langle\dot{M}\boldsymbol{v}'\rangle = n_s h_c\langle\dot{M}\rangle\boldsymbol{u}$ in the momentum equation, we obtain equation (17d) in the invariant vector form

$$\nabla(\tfrac{1}{2}u^2) + (\nabla\times\boldsymbol{u})\times\boldsymbol{u} = -\nabla(\mathcal{U} - \tfrac{1}{2}\Omega_p^2\varpi^2) - 2\boldsymbol{\Omega}_p\times\boldsymbol{u} - \rho^{-1}\nabla P\,. \tag{25}$$

[5] The difference between the fluid velocities is probably less than that estimated on the basis of hydrodynamic drag alone since the lines of force of the interstellar magnetic field which thread through both phases probably provide significant coupling of their motions. An estimate based on the square of the Alfvén speed $H^2/4\pi[F_c\langle\rho_c\rangle + (1 - F_c)\rho]$ suggests that the presence of the magnetic field would also hinder slightly the compression of the general medium (cf. Roberts and Yuan 1970).

The jump conditions (18a), (18b), and (18d) now become

$$[f_c\langle\rho_c\rangle u_\perp]_1^2 = \delta_m\,, \quad \delta_m \equiv \nabla_m/(1 - F_c)h\,, \tag{26a}$$

$$[\rho u_\perp]_1^2 = -\delta_m\,, \tag{26b}$$

$$[f_c\langle\rho_c\rangle(u_\perp^2 + c^2) + \rho u_\perp^2 + P]_1^2 = 0\,. \tag{26c}$$

In equation (26c) we have eliminated Δ_p by adding equations (18c) and (18d).

Given ζ, c, $\mathcal{U}$, together with the subsidiary equations (7), (14), and the jump conditions (20), (21), and (26), we may solve equations (25), (24), and (15), or equations (25), (22a), and (22b) to obtain the five variables ρ, $\boldsymbol{u}$, f_c, and $\langle\rho_c\rangle$. Equations (16) and (23) may then be used to recover h, h_c, and F_c.

Equations (22), (24), and (25) together with the scaling laws discussed in § II*c* allow a simple transformation. *If ζ is changed by a multiplicative constant λ, we obtain the new solution simply by multiplying $\langle\rho_c\rangle$, ρ, and P by the same factor λ, but leaving f_c, h, $\boldsymbol{u}$, and the position in the Galaxy unchanged.*

f) Asymptotic Form of the Equations

We suppose that the gravitational potential $\mathcal{U}$ can be decomposed into a static axisymmetric potential $\mathcal{U}_0$ and a spiral potential $\mathcal{U}_1$ which is static when written in rotating cylindrical coordinates $(\varpi, \phi = \theta - \Omega_p t, z)$. Similarly, we write $\boldsymbol{u} = \boldsymbol{u}_0 + \boldsymbol{u}_1$, $\rho = \rho_0 + \rho_1$, where $\boldsymbol{u}_0$ is a reference flow chosen to provide pure centrifugal balance against the radial field in the plane $z = 0$:

$$-\nabla\mathcal{U}_0(\varpi) \equiv -\boldsymbol{e}_\varpi\varpi\Omega^2(\varpi)\,, \qquad \boldsymbol{u}_0 = \boldsymbol{e}_\phi\varpi(\Omega - \Omega_p)\,, \tag{27}$$

and where ρ_0 is a reference mass distribution for the intercloud gas chosen to be axisymmetric and (nearly) equal to the density of the intercloud medium averaged along a circle of radius ϖ.

We follow Roberts (1969) and introduce now the orthogonal curvilinear coordinates (η, ξ, z) where η = constant defines a curve of constant phase of the spiral gravitational potential (cf. eq. [29] with $\sin[i] \ll 1$). For a small range of radii in the neighborhood of a nearly circular streamline, we may approximate the geometric form of the spiral by a logarithmic spiral. In this approximation, the coordinates (η, ξ) and the associated unit vectors $(\boldsymbol{e}_\eta, \boldsymbol{e}_\xi)$ are obtained from (ϖ, ϕ) and $(\boldsymbol{e}_\varpi, \boldsymbol{e}_\phi)$ by the transformation

$$\begin{pmatrix}\eta\\ \xi\end{pmatrix} = \begin{pmatrix}\cos i & \sin i\\ -\sin i & \cos i\end{pmatrix}\begin{pmatrix}\ln\varpi\\ \phi\end{pmatrix}, \quad \begin{pmatrix}\boldsymbol{e}_\eta\\ \boldsymbol{e}_\xi\end{pmatrix} = \begin{pmatrix}\cos i & \sin i\\ -\sin i & \cos i\end{pmatrix}\begin{pmatrix}\boldsymbol{e}_\varpi\\ \boldsymbol{e}_\phi\end{pmatrix}. \tag{28}$$

The spiral gravitational field (in the plane $z = 0$) associated with $\mathcal{U}_1$ is of the form

$$-\nabla\mathcal{U}_1 = \boldsymbol{e}_\eta g_1 = -\boldsymbol{e}_\eta F\varpi\Omega^2(\varpi)\sin(-2\eta/\sin i)\,, \tag{29}$$

where F is the strength of the spiral field expressed as a fraction of the local axisymmetric field.

If we assume that $\sin i \ll 1$, we may derive the nonlinear asymptotic form of the fluid equations in a way similar to that given by Roberts. *Under the assumption that ζ is a constant along a streamline, the results along the part of the streamline which contains no phase transitions can be written*

$$\rho u_\eta = \text{constant}\,, \quad f_c\langle\rho_c\rangle u_\eta = \text{constant}\,, \tag{30a}$$

$$\frac{\partial u_{n1}}{\partial\eta} = (u_{\eta 0} + u_{\eta 1})\frac{(2\varpi\Omega u_{\xi 1} + \varpi g_1)}{(u_{\eta 0} + u_{\eta 1})^2 - a^2}\,, \tag{30b}$$

$$\frac{\partial u_{\xi 1}}{\partial\eta} = -\frac{\varpi\kappa^2}{2\Omega}\frac{u_{\eta 1}}{(u_{\eta 0} + u_{\eta 1})}\,. \tag{30c}$$

In equation (30b) a^2 is the square of the effective speed of sound for large-scale pressure disturbances in the intercloud medium and is defined by

$$a^2 = \left(\frac{\partial P_{eq}}{\partial \rho}\right)_\zeta . \tag{31}$$

The quantities $u_{\eta 0}$ and $u_{\xi 0}$ (denoted, respectively, as $w_{\perp 0}$ and $w_{\parallel 0}$ by Roberts 1969) are the η and ξ components of the reference velocity while κ^2 is the square of the epicyclic frequency:

$$u_{\eta 0} = \varpi(\Omega - \Omega_p) \sin i , \qquad u_{\xi 0} = \varpi(\Omega - \Omega_p) \cos i , \tag{32a}$$

$$\kappa^2 = \frac{2\Omega}{\varpi} \frac{d}{d\varpi} (\varpi^2 \Omega) . \tag{32b}$$

Along that part of the streamline which does contain phase transition, the fluid equations read

$$(f_c\langle \rho_c \rangle + \rho) u_\eta = \text{constant} , \tag{33a}$$

$$\frac{\partial u_{\eta 1}}{\partial \eta} = \frac{2\varpi\Omega u_{\xi 1} + \varpi g_1}{(u_{\eta 0} + u_{\eta 1})} , \tag{33b}$$

$$\frac{\partial u_{\xi 1}}{\partial \eta} = - \frac{\varpi\kappa^2}{2\Omega} \frac{u_{\eta 1}}{(u_{\eta 0} + u_{\eta 1})} . \tag{33c}$$

Once the velocities $u_\eta = u_{\eta 0} + u_{\eta 1}$ and $u_\xi = u_{\xi 0} + u_{\xi 1}$ have been obtained from an integration of equations (30) or (33), the position in the galaxy (η, ξ) may be recovered from the characteristic equation

$$\frac{d\xi}{d\eta} = \frac{u_\xi}{u_\eta} . \tag{34}$$

We conclude this section with the following comment. Comparison of the jump conditions which can be derived from equations (30) and (33) with those of equations (26) reveals that a shock is consistent with the asymptotic approximation used above only if $\boldsymbol{e}_\eta$ is the direction normal to the shock front. *Hence, we have the important consistency requirement that the shock front must lie asymptotically close to a curve of constant phase of the spiral gravitational potential.* Consequently, the justification for the asymptotic approximations taken here can come only *a posteriori*. We need to obtain the family of STS solutions (*streamtube* band through *two* periodically located *shocks*) at different radii to construct the shock front in the Galaxy which forms the TASS pattern (*two-armed spiral shock* pattern). If the TASS pattern does not pass the consistency test, a two-dimensional integration of equations (22), (24), and (25) would be required.

IV. THE LARGE-SCALE GALACTIC FLOW

The asymptotic equations which were derived in § III*f* can be integrated numerically to determine the steady large-scale flow of gas in the Galaxy. We are primarily interested in doubly periodic solutions which contain shocks.[6] Numerical calculations have been

[6] Nonlinear solutions without shocks are also possible (Vandervoort 1971). It is not possible to state at the present time the regimes of applicability of the two types of solutions. We feel solutions with shocks to be natural candidates whenever the spiral gravitational field is sufficiently strong to force the flow along a closed streamline to be subsonic in part and supersonic elsewhere. The terminology "subsonic," "sonic," and "supersonic" is used here to refer to the magnitude of u_η compared with that of a. The shocks are actually highly oblique, and the flow is always hypersonic if the total velocity $u = (u_\eta^2 + u_\xi^2)^{1/2}$ is used instead of the normal component u_η to define the Mach number. See Appendix D.

carried out for regions lying at distances 4–12 kpc from the galactic center. Because the qualitative features are similar throughout this region, we shall present only the solutions for the streamline which passes through the solar neighborhood. It has been verified, by the way, that the TASS pattern in the range 4–12 kpc does in fact satisfy the consistency requirement mentioned at the end of § III*f*.

a) Uniqueness of the STS Solution

Given an equilibrium model for the Galaxy and the value of the pattern speed Ω_p, the properties of the underlying spiral gravitational field are completely determined up to an arbitrary multiplicative constant for the relative strength F (Shu, Stachnik, and Yost 1971). In our calculations we adopt a mass model which is very similar to Schmidt's (1965) model and a pattern speed $\Omega_p = 13.5$ km s^{-1} kpc^{-1} (Lin *et al.* 1969). For $\varpi = 10$ kpc, the circular frequency $\Omega = 25$ km s^{-1} kpc^{-1}, the epicyclic frequency $\kappa = 31$ km s^{-1} kpc^{-1}, the z-oscillation frequency $(\partial^2\mathcal{U}/\partial z^2)^{1/2} \simeq (\partial^2\mathcal{U}_0/\partial z^2)^{1/2} = 86$ km s^{-1} kpc^{-1}, while the pitch angle i of the spiral is 6°.9. We arbitrarily assume c to be a constant, 8 km s^{-1}.

In Appendix C we present the details of the procedure used to solve equations (30) and (33). Here, we need only remark that the character of the STS solution is uniquely specified once we are given F_c, the fraction of volume occupied by the clouds, at *one* point along the streamline (or alternatively, once we are given the average vlaue of F_c along the streamtube). In principle, this value can be determined by knowledge of the total amount of gas present locally in the Galaxy. In practice, it is a free parameter which can be adjusted to give a best fit for observations relating to the emission and absorption at 21 cm.

For given F_c small compared with unity, the solutions fall into three general categories if we vary the value adopted for F, the relative strength of the spiral gravitational field. With $F < 0.010$ (at $\varpi = 10$ kpc), no solutions with two periodically located shocks can be found; presumably, the flow occurs in this case without developing shocks. With $0.010 < F < 0.033$, solutions containing shocks are found, but phase transitions do not occur anywhere along the streamline (in a steady state) because the total variation of the intercloud pressure does not span $(P_{\min}, P_{\max})$. With $F > 0.033$, phase transitions are unavoidable if $F_c \neq 0$. We present below the results of the calculations for the cases $F = 0.033$, $F_c = 0$ everywhere along the streamline (no clouds at all) and $F = 0.050$, $F_c \neq 0$ anywhere along the streamline. The value $F = 0.050$ corresponds to the field strength believed on the basis of other dynamical studies to be actually present in the solar neighborhood (Yuan 1969).

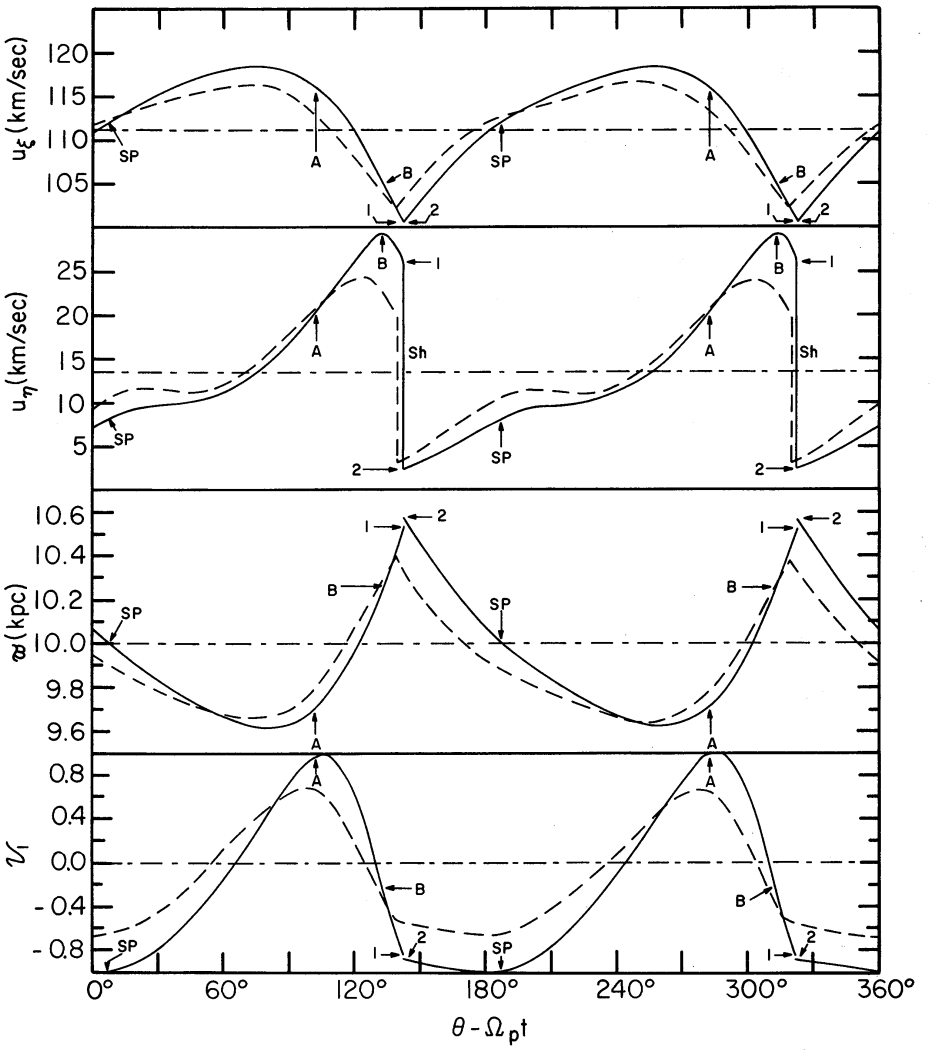
b) Results

Figures 2*a* and 2*b* show the results of the calculations for the galactic flow when $F = 0.033$, $F_c = 0$ (*dashed lines*) and when $F = 0.050$, $F_c \neq 0$. The latter case is computed with F_c chosen so that the space density of atomic hydrogen is 0.5 cm^{-3} when averaged over the streamtube. The clouds then formally constitute 94 percent of the gaseous matter immediately after the galactic shock; however, this relatively large percentage may be reduced if the layer of intercloud gas is, as is likely, considerably thicker than that given by the thermal scale height. In any case, the dynamics of the intercloud medium is insensitive to the exact fraction of mass contained in the cloud phase.

The reference axes in Figures 2*a* and 2*b* have been chosen so that $\theta - \Omega_p t = 0°$ and 180° are the locations of the minima of the spiral gravitational potential (normalized in the figures to -1 for $F = 0.050$) as encountered by the gas in its flow around the Galaxy. The displacement of the shock from the location of the potential minimum is ~37° in $\theta - \Omega_p t$ but is only ~12° in the half-phase $\eta/\sin i$ of $\mathcal{U}_i$. (The variable $\eta/\sin i$ is unfortunately denoted as θ_{Ω_p} by Roberts 1969.)

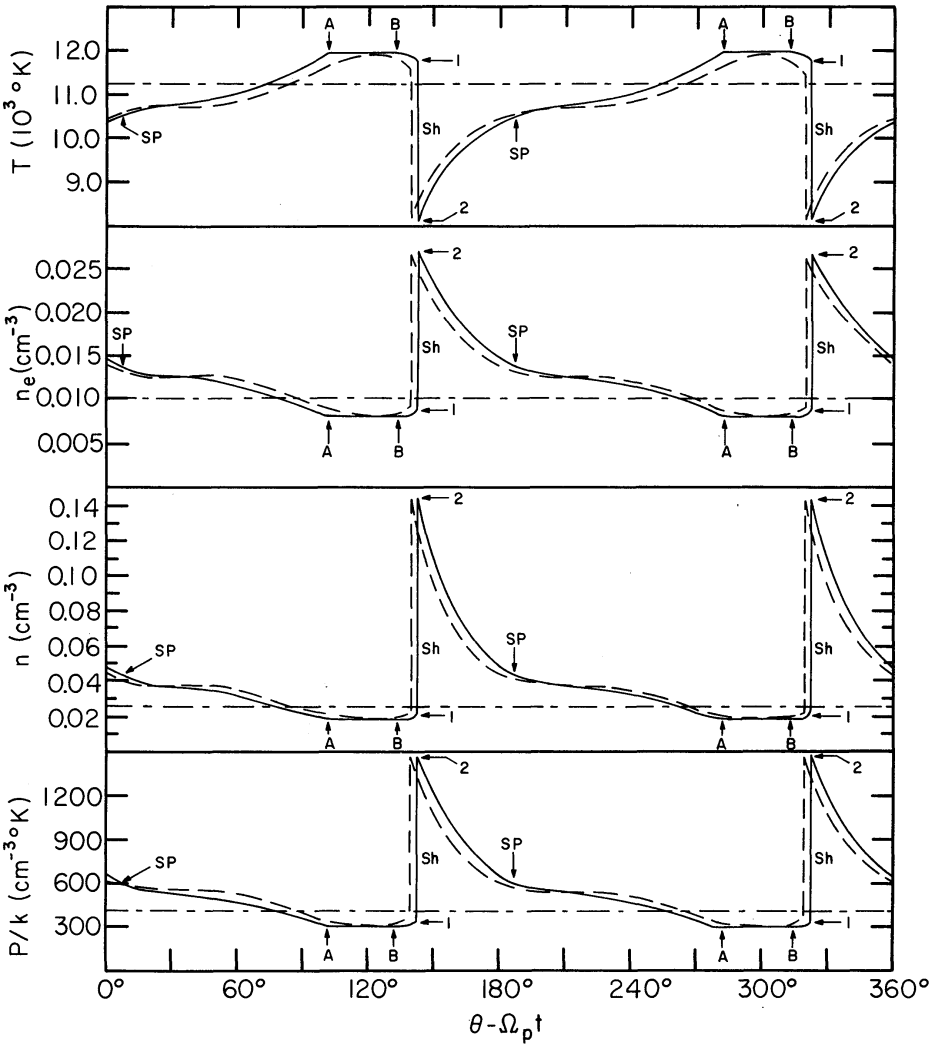
The dynamical solution for the case $F = 0.033$, $F_c = 0$ bears a great similarity to that calculated by Roberts (1969) on the basis of a one-component "turbulent" model of the interstellar medium. However, our results refer to the intercloud medium rather than to the clouds, and the sound speed a is associated with the gas kinetic temperature of this medium rather than with the velocity c of random cloud motions. Nevertheless, because c is numerically close to a and because the compression of the intercloud medium occurs nearly isothermally, the large-scale dynamics is similar for the two models.

The results for $F = 0.050$, $F_c \neq 0$, present a more visible difference between the present calculations and previous ones in the inclusion of the effects of phase transitions. Notice that the jumps in pressure and density across the shock are nearly the same for the solid and dashed curves but that the jump in the normal component of the velocity



(a)

FIG. 2.—The variation of the properties of the intercloud gas along a streamline which passes through the solar neighborhood. The solid curves correspond to $F = 0.050$; the dashed curves, to $F = 0.033$. For the case $F = 0.050$, the horizontal dashed-dotted lines give the "average" values of the various quantities, whereas the labels SP, A, B, 1, Sh, 2 correspond to special points whose properties are discussed in the text. (*a*) The slight nonclosure seen in the curves for the radius ϖ of the streamline is discussed in § IV*c*. (*b*) Our definitions of the "averages" of the pressure, the atomic density, the electron density, and the temperature give more weight to the regions of decompression (e.g., between points A and B) where the gas occupies greater volume (see Appendix C).



(b)

FIG. 2.—*Continued*

is substantially larger for the solid curve. The excess of postshock pressure which should have resulted for the solid curve is "used" to transform intercloud material into cloud material.

To discuss the solution for the solid curve in more detail, let us follow the flow along the streamline. Immediately after the shock (labeled 2) the intercloud pressure is, of course, very large—indeed, $P(2) \simeq P_{\max}$ in this case. Despite the hindrance of the perturbation Coriolis force, the perpendicular component u_η of the gas velocity increases as the gas accelerates under the combined action of the pressure and gravitational forces. This acceleration continues and u_η reaches the speed of sound a. The requirement that the flow pass smoothly through the sonic point (labeled SP) coupled with the condition that the flow be doubly periodic determines uniquely the location of the shocks in the Galaxy (see Appendix C for details).

After passage through the sonic point, the gas expands into supersonic flow, and the acceleration caused by the Coriolis force eventually overcomes the deceleration caused by the potential well. As the gas pressure drops steadily, it reaches the value $P_{\min}$ at a

point (labeled A) where the onset of thermal instability begins to transform cloud material into intercloud material. At point A we shift our integration from equations (30) to equations (33). The flow proceeds with phase transition maintaining the pressure at $P_{\min}$ until the point (labeled B) where u_η reaches a maximum value and starts to decrease. The ensuing flow is compressional, and the transformation of cloud material into intercloud material ceases. Thus, at point B we shift our integration from equations (33) back to equations (30). The flow now proceeds without phase transition until the gas shocks by slamming into gas traveling at subsonic speeds halfway around the Galaxy from where we first picked up the flow. Inside the shock layer (labeled Sh) there is a transformation of intercloud material into cloud material which is exactly equal to the total amount of the reverse transformation during the flow from point A to point B. In the case shown, $\sim$31 percent of the local intercloud material at point 1 becomes cloud material by the time point 2 is reached; however, this number can be substantially increased (indirectly) if depletion of cooling agents (notably C^+) onto grains occurs in clouds and effectively raises the value of $P_{\min}$ (Field, private communication).

c) Nonclosure of the Streamlines

The solutions are doubly periodic in every variable presented in Figures 2*a* and 2*b* *except* the instantaneous radius ϖ of the streamline and the value of the spiral gravitational potential $\mathcal{U}_1$ (which depends on ϖ) seen by the gas in its flow. After a 180° circuit around the Galaxy, the gas is actually found at a slightly smaller radius than that at which it began its motion.

It is difficult to separate the fraction of the effect that is physical and due to the torque exerted on the gas by the underlying spiral gravitational field (cf. the criticism of Roberts's original calculations by Kalnajs as quoted by Simonson 1970) from the fraction that is mathematical and due to the approximate nature of the asymptotics adopted for our calculations. After all, the reduction in § III*f* of a two-dimensional problem to a one-dimensional integration in the variable η entails a loss of one degree of freedom which leaves us without any control over the radial variation of the streamlines.

Regarded in the above mathematical sense, the nonclosure of the streamline ($\sim$5 percent of the *variation* in ϖ) is not serious since it is asymptotically small (of order $\sin i$). The deeper physical aspect of the irreversibility associated with shocks is discussed briefly in footnote 4 and in § VII.

d) The Experience of the Clouds in the Galaxy

Figure 3 presents the response of the clouds to the periodic compressions and decompressions of the intercloud medium for the case $F = 0.050$. The two curves for F_c are explained in the figure caption. The variation of the cloud radius R_c is computed for a 400 $M_\odot$ cloud without including the effects of self-gravitation.

The most important feature of figure 3 is the following. Even under the assumption that ζ is uniform, the interstellar clouds show highly nonuniform *mean* properties because of the substantial variations of pressure induced by the underlying spiral gravitational field. Figure 3 shows that along a given streamline, the clouds' atomic density $\langle n_c \rangle$, temperature $\langle T_c \rangle$, electron density $\langle n_{ec} \rangle$, and "filling fraction" F_c may vary, respectively, by a factor 40, 8, 4, and 5. These substantial variations are *systematic* and provide, therefore, an excellent basis to test the theory. In reading these graphs, it should be borne in mind that the velocity interval spanned by points A and B may well constitute most of what radio astronomers generally attribute to the "interarm region." Individual clouds in the maximally decompressed state would be very difficult to detect in 21-cm absorption. Indeed, even if only thermal broadening is considered, we estimate a peak optical depth $\sim$0.05 to result for a line of sight passing through the center of a spherical 400 $M_\odot$ cloud which is in the fully decompressed state.

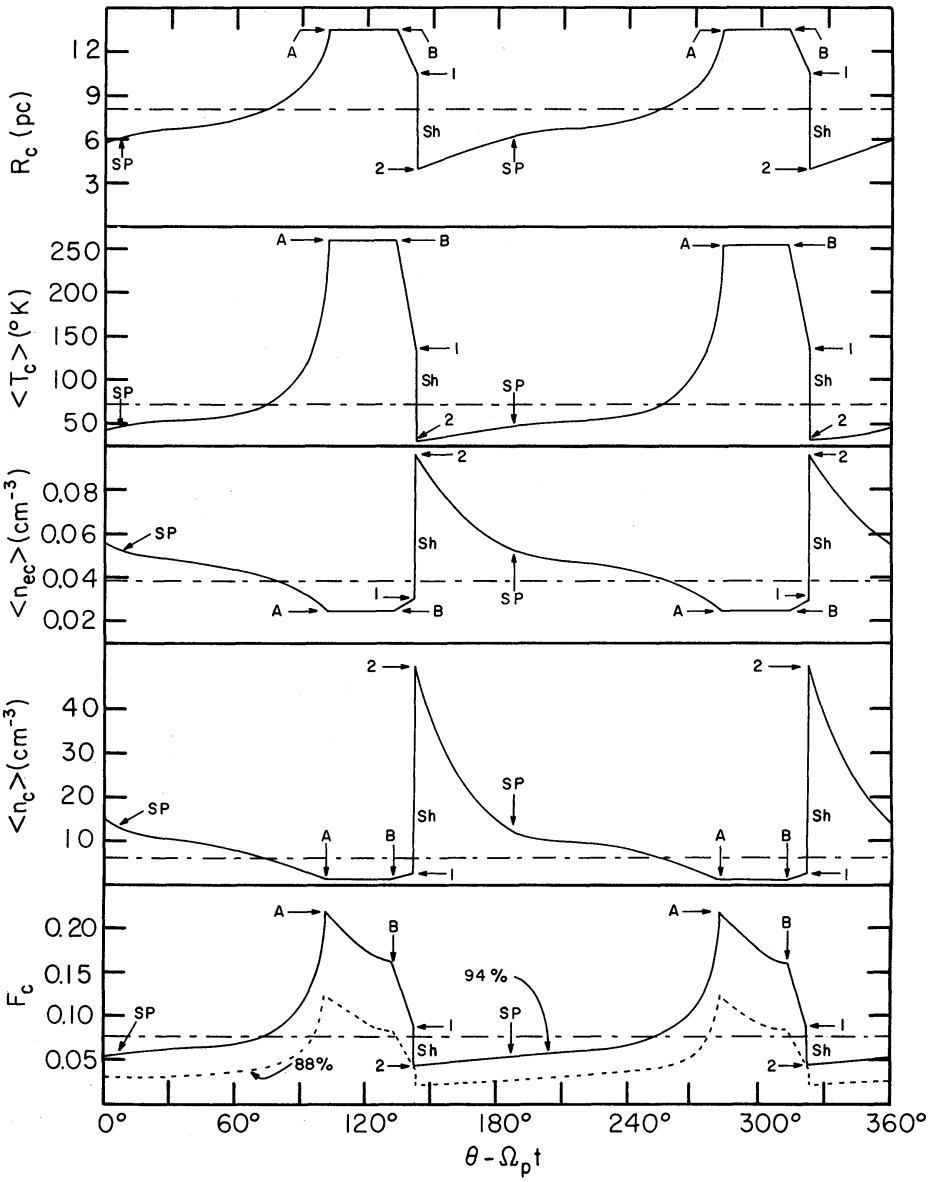


FIG. 3.—The variation of the properties of the cloud phase along a streamline which passes through the solar neighborhood. The solid curves correspond to the case $F = 0.050$ with F_c chosen so that the clouds constitute 94 percent of the gaseous matter immediately after a galactic shock (at point 2). The dotted curve for F_c is computed for the same field strength $F = 0.050$ but under the assumption that the clouds constitute 88 percent of the gaseous matter immediately after a galactic shock. The variations of the other properties of the cloud phase are nearly indistinguishable for the two cases.

e) Some Preliminary Remarks on 21-Centimeter Absorption Measurements

If the systematic departures of the clouds' mean motions from pure (differential) rotation are ignored, we easily calculate that peak optical depths as large as ~7 can be reached in spiral arms, whereas optical depths as low as ~0.07 can be reached in regions between spiral arms. These correspond roughly to the observed values, although some workers have attributed much smaller optical depths to "interarm regions" (cf. Quirk 1971).

The interpretation of the combined emission and absorption profiles in the direction of *Galactic* continuum sources requires special care. It is necessary to take into account the systematic motions along the line of sight as well as the variations of the internal properties of the clouds. In an oft-quoted case—in the direction of Cas A (de Jager 1970)—we discovered upon detailed examination that the velocity interval which is generally attributed to the "interarm" region may correspond, in fact, to *a small part of the shock layer!* (Because the velocity changes very rapidly in the shock layer, a given velocity *interval* corresponds to a small part of the thin shock layer. We shall henceforth refer to this effect as "velocity stretching.") Furthermore, a large region well behind the Perseus "spiral arm" possesses a small gradient of the line-of-sight velocity ("velocity crowding") and may contribute to the emission seen at the original velocity interval because of some spillover due to the velocity dispersion of the clouds. If Cas A lies between the shock layer and the region of "velocity crowding" (i.e., somewhere in the Perseus spiral arm), as seems likely, then the observed "anomalous" combination $\tau \lesssim 0.02$ and $T_b \sim 15°$ K can be understood because the (negligible) absorption is due to a small part of a thin shock layer which is quite distinct from the region of "velocity crowding" that gives rise to the emission. We are now in the process of making detailed velocity-profile calculations to verify these conjectures.

V. THE INTERNAL STRUCTURE OF A SHOCK LAYER

In the previous sections we have treated the shock layer as a discontinuity in the flow. In this section we calculate the internal structure of the shock layer in the intercloud medium. As already mentioned in § I, we do not consider here any of the interactions (e.g., phase transitions and drag) with the clouds. Thus, strictly speaking, our calculations apply only if F_c is identically zero; nevertheless, in § V*d*, we shall make some comments on the behavior of the clouds when they are treated as test particles injected into the shock layer.

Whenever the intercloud gas is out of equilibrium, we can no longer bypass the energy equation, the charge-balance equation, and the equation of state, and simply use the equilibrium P-ρ relation. To make this discussion more concrete we follow Whitney and Skalfuris (1963) (see also Field *et al.* 1968) in distinguishing four possible regions in the vicinity of the shock layer: (i) a "preshock" region which may be influenced by the radiation emanating from the shock, (ii) a "viscous" region where elastic scattering redirects much of the normal component of the bulk momentum, (ii) an "internal relaxation" region where rapid ionizations and excitations occur to adjust the level populations to local statistical equilibrium, and (iv) a radiative region where the shock-deposited energy is emitted as radiation and the gas cools to a state of local thermal equilibrium consistent with the downstream conditions.

In our calculations for the streamline which passes through the solar neighborhood, the shock speeds obtained are of the order of 20–30 km s^{-1}. Since the incoming velocity does not exceed $\sim$50 km s^{-1}, corresponding to $\sim$1 rydberg, the wavelengths of the radiation emitted in the shock layer lie longward of the Lyman continuum, and the radiative precursor can be ignored.

The thickness of the viscous layer is a few atomic mean free paths, $\sim 10^{-2}$ pc. Thus, region ii—and, for that matter, region iii—is very thin in comparison with the entire relaxation layer. We shall not discuss the well-known techniques (see, e.g., Zel'dovich and Raizer 1966) for finding the internal structure of the "microscopic relaxation" layers ii and iii but shall treat them as discontinuities in the flow.

Recombination and radiative cooling have the longest characteristic time scales, and it is the structure of region iv which we wish to consider in detail. Under the assumption that local statistical equilibrium prevails and that the medium is optically thin to the radiation emitted in the shock layer, the energy equation, the charge-balance equa-

tion, and the equation of state for the intercloud medium read

$$\nabla\cdot(\epsilon u) + P\nabla\cdot u = \mathcal{H}(n, n_e, T, \zeta) , \tag{35a}$$

$$\nabla\cdot(n_e u) = \mathcal{I}(n, n_e, T, \zeta) , \tag{35b}$$

$$P = (n + n_e)kT . \tag{35c}$$

In the above the internal energy per unit volume ϵ is the microscopic kinetic energy plus the ionization energy:

$$\epsilon = 3P/2 + n_e\chi , \tag{36}$$

where χ is the mean ionization potential of all atoms and ions whose ionization states can be changed inside the shock layer. To complete the set of equations we require equations (2), (22b), and (25).

Because we shall not discuss the internal structure of the viscous layer, we have not included the viscous stress term in the momentum equation (22b) nor have we included the viscous-dissipation and heat-conduction terms in the energy equation (35a). Thus, we may need to supplement the fluid equations by jump conditions of the Hugoniot type.

In the general galactic flow, the various gradients are small, and the left-hand terms of equation (35a) are easily verified to be at least one order of magnitude smaller than either Γ or Λ (see § II*a* for the definitions of Γ and Λ). A similar statement holds for equation (35b). Thus, outside the shock layer, it is a good approximation to take $\Gamma = \Lambda$, and $I = R$. These relations are, of course, equivalent to the assumption (1) of local thermal equilibrium.

Inside the shock layer, the gradients of $u_\perp = u_\eta$, P, and ρ are large, and it is no longer correct to set the right-hand sides of equations (35a) and (35b) equal to zero. We assume, however, that the shock layer is thin so that the variations of $\mathcal{U} - \frac{1}{2}\Omega_p^2\varpi^2$ and $u_\parallel = u_\xi$ can be ignored inside the shock layer. Consistent with this approximation, we may integrate equations (22b) and (25) in the direction normal to the shock front to obtain

$$\rho u_\eta = \text{constant} \equiv j_0 , \tag{37a}$$

$$P + \rho u_\eta^2 = \text{constant} \equiv p_0 . \tag{37b}$$

If we write $ds = \varpi d\eta$, equations (35a) and (35b) can now be written in the equivalent one-dimensional form

$$(5P/2 - 3\rho u_\eta^2/2)\frac{du_\eta}{ds} = (5p_0/2 - 4j_0u_\eta)\frac{du_\eta}{ds} = \mathcal{H} - \mathcal{I}\chi , \tag{38a}$$

$$\frac{d}{ds}(n_e\chi) = \mathcal{I} . \tag{38b}$$

We have transformed the energy equation to the more useful form (38a) by using equations (36) and (37) to eliminate all derivatives other than du_η/ds; furthermore, we have assumed that χ is a constant and have used equation (38b) to rewrite $d(n_e\chi u_\eta)/ds$ as $\mathcal{I}\chi$. The term $5P/2$ in equation (38a) corresponds, of course, to the enthalpy per unit volume of a perfect monatomic gas.

We now show that a viscous layer must precede any other relaxation phenomenon. From equation (38a), we see that the gradient of the velocity formally becomes infinitely steep if u_η tries to pass the value $5p_0/8j_0$. This happens when u_η is equal to the local *adiabatic* speed of sound $a_s = (5P/3\rho)^{1/2}$.[7] The formal occurrence of large velocity

[7] The adiabatic speed of sound a_s is always larger than the effective speed of sound a for large-scale pressure disturbances defined through equation (31). If the incoming velocity is subsonic with respect to a_s but supersonic with respect to a, there can in principle be a weak thermal shock which is not preceded by a viscous layer. Such situations do not usually arise in practice.

gradients indicates, of course, that viscosity cannot be neglected. A viscous layer occurs physically because the incoming gas is incident "deafly" upon gas immediately downstream from it, and the slamming of gas into gas produces a thin viscous shock.

a) Jump Conditions for the Viscous Layer

We treat the viscous shock as a discontinuity and obtain the jump conditions on energy and electron density by integration of equations (38) across the viscous layer when the right-hand sides are taken to be continuous:

$$[(5p_0/2 - 2j_0u_\eta)u_\eta]_1^{1'} = [(5P/2 + \tfrac{1}{2}\rho u_\eta^2)u_\eta]_1^{1'} = 0\,, \tag{39a}$$

$$[n_e u_\eta]_1^{1'} = 0\,. \tag{39b}$$

The indices 1 and 1′ refer to positions immediately preceding and immediately following the viscous shock layer. Equations (39a), (37a), and (37b) lead, of course, to the usual Hugoniot relations while equation (39b) together with equation (37a) expresses the simple fact that the fractional degree of ionization is "frozen" in the viscous layer. The jump condition on the temperature is, of course, obtained from the equation of state (35c).

b) Method of Solution and Some General Considerations

The mathematical prescription for obtaining the structure of the complete shock layer is now clear. We begin by subjecting the incoming flow to a viscous shock. With the shocked "initial conditions at 1′," we integrate equations (38a) and (38b) together with the subsidiary equations (2), (35c), (37a) and (37b). The critical velocity $5p_0/8j_0$ is always spanned inside the viscous layer so there are no further difficulties involving large velocity gradients.

In general the viscous shock throws the gas badly out of equilibrium so that the right-hand sides of equations (38a) and (38b) are no longer zero. The integration is carried as far as is needed for the variables to relax asymptotically to the values appropriate for the far-downstream equilibrium conditions. This will eventually occur since the right-hand sides of equations (38a) and (38b) force the variables to change until $\mathcal{H}$ and $\mathcal{G}$ are both zero. It is easy to verify that the far-downstream conditions will automatically be consistent with those obtained at point 2 from the jump conditions (26) when δ_m and f_c are taken to be zero. Hence, what appears at first sight to be a boundary-value problem reduces, in fact, to a much simpler initial-value problem.

If we plot the "shock adiabat" in the usual way, we find that there is, of course, an increase in the specific entropy of the gas as it crosses the viscous layer (from point 1 to point 1′); but there is, in fact, a net *decrease* from point 1 to point 2. This last result can be seen from the approximate polytropic relation (6) which relates the equilibrium states 1 and 2 with an effective $\gamma = (N + 1)/N \simeq 0.75 < 5/3$ for the intercloud gas. The net decrease of entropy does not violate the second law of thermodynamics inasmuch as the gaseous system considered independently of the emitted radiation field and the external sources of heating does not constitute a closed system.

There exists a simple scaling law which applies to the structure of the shock layer. If we multiply ζ by λ, the mass flux j_0 and momentum flux p_0 (determined by the calculation of § IV) also change by a factor λ. Therefore, the solution for equations (37) and (38) can be obtained simply by scaling $\rho = mn$, P, n_e, *and* s^{-1} by the same factor λ, but leaving u_η and T unchanged. The inverse scaling of s (i.e., the thickness of the relaxation layer) by the factor λ is necessary because of the equidimensional property (4) of the functions $\mathcal{H}$ and $\mathcal{G}$.

c) Results

Figure 4 shows the interior solution for the shock in the critical case when the strength of the spiral gravitational field is 3.3 percent of the axisymmetric field. The abscissae

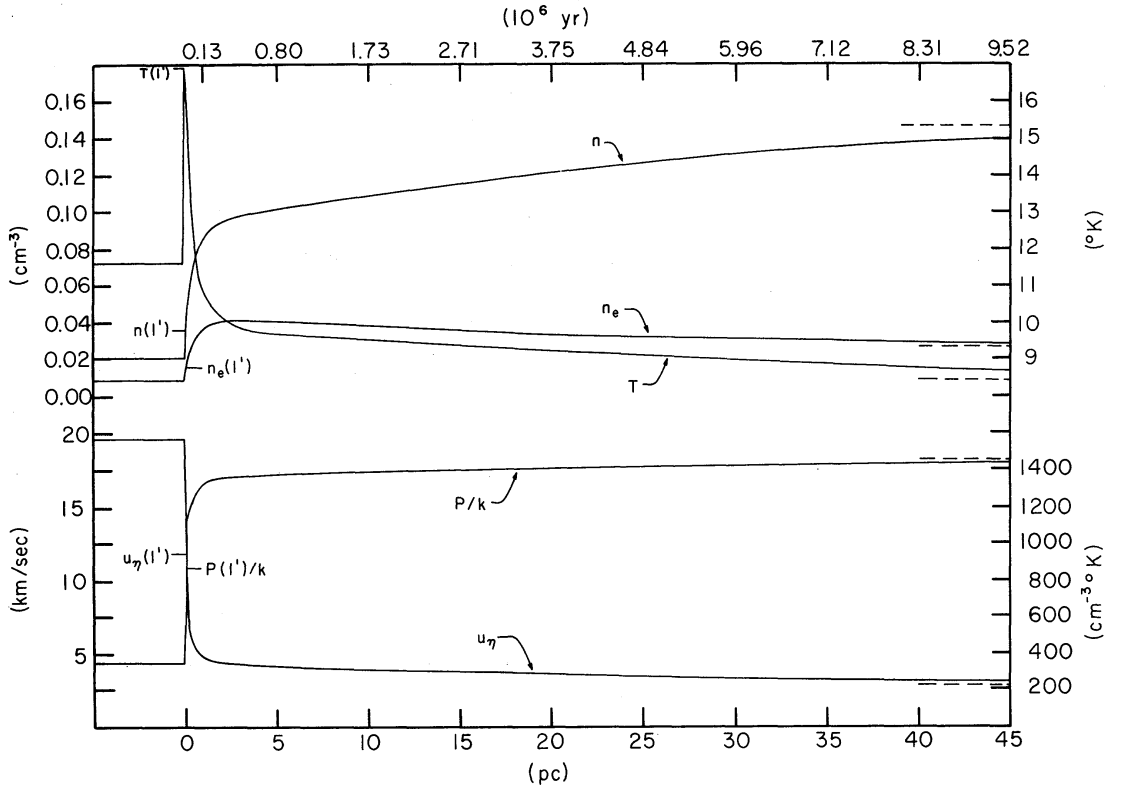


FIG. 4.—The internal structure of the shock layer in the intercloud gas when all interactions with the cloud medium are ignored. The preshock (*horizontal solid lines*) and postshock (*horizontal dashed lines*) conditions correspond to those obtained in the solution for the large-scale galactic flow when the strength of the spiral gravitational field is taken to be 3.3 percent that of the local axisymmetric field. Immediately after the "viscous layer," the velocity, pressure, atomic density, electron density, and temperature attain the values marked by the horizontal ticks.

show the distance and time elapsed as we follow the flow downstream in the direction normal to the shock front. The horizontal solid lines indicate the incoming upstream values; the horizontal dashed lines, the asymptotic downstream values. The initial jumps in the viscous layer are marked by the ticks.

We call attention to the particularly fast relaxation for the pressure. The intercloud pressure reaches 90 percent of its downstream value in less than 1 pc, i.e., less than 0.13×10^6 years after the beginning of the shock. Note also the relaxation phenomenon for the thermal variables T and n_e. There is a tremendous transient heating in the viscous layer, followed by rapid radiative cooling which results in a temperature "spike." Thereafter, T relaxes asymptotically in a long tail to its downstream equilibrium value. The fractional ionization is frozen in the viscous layer, and the density of free electrons suffer initially a simple compression. Thereafter, the increased temperature results in a higher rate of ionizations by electron impact and a lower rate of recombinations by electron capture, and n_e overshoots its downstream asymptotic value. Eventually, and inevitably, recombinations reassert themselves and n_e settles asymptotically to the downstream value. We note that the overshoot of T and n_e is not depicted in the "exterior solution," figure 2*b*, for the galactic shock because it is hidden as detail of a region approximated to be infinitesimally thin. Furthermore, consistent with the assumption that the case considered lacks phase transition, no overshoot occurs in figure 4 for the pressure P.

The width of the entire shock layer, ~50 pc, is determined by the slowest relaxation process—in this case, the process of recombination. The time 1.0×10^7 years spent in the shock layer is short in comparison with the time 2.7×10^8 years spent between shocks, so the treatment of the entire shock layer as a discontinuity in the large-scale flow is well justified. The situation may be less favorable, however, if interactions with the clouds are taken into account.

d) Comments on the Behavior of the Clouds in the Shock Layer

Figure 4 shows the normal component u_η of the intercloud gas to be sharply decelerated in the galactic shock layer. The deceleration of the clouds must be much less abrupt since it arises as a result of the drag between the two phases and the diffusive transport of momentum due to cloud-cloud collisions. Thus, a test cloud entering the shock layer at the upstream fluid velocity ~20 km s^{-1} would cross one cloud diameter, say 20 pc, in 1×10^6 years and would then find itself in a region where the gas kinetic pressure is ~5 times higher than its internal pressure (which has had no time to adjust to the new conditions) and where the velocity differential between it and the ambient medium is ~15 km s^{-1}. The resulting effects on the cloud are too complicated to discuss in detail here, but we give below the general qualitative features which can be expected to be present.

The sudden pressure loading would send a shock, traveling at 2–3 km s^{-1}, into the interior of the cloud. This shock would probably gain in strength as it propagates inward, because of the implosion geometry. The internal shock is the mechanism by which the cloud is compressed to its downstream conditions. Despite the nearly simultaneous initiation of the pressure pulse on the entire surface of the cloud, the internal shock front would lack spherical symmetry because the "intercloud wind" resulting from the relative motion tends to produce a highly asymmetric distribution of dynamic pressure which is largest at the leading face of the cloud. This would tend to flatten the cloud somewhat. (At low relative speeds and comparable gas kinetic pressures inside and outside the cloud, the asymmetry in the distribution of dynamic pressure would tend to drive a dipole-like internal circulation in the cloud which offsets the tendency for the cloud to flatten.)

If we write w for the normal component of the cloud's velocity relative to the intercloud medium, A for its frontal area, and l for its typical thickness in the direction of travel, the equation of motion for the test cloud can be written

$$\rho_c A l \frac{dw}{dt} = -CAw^2/2 , \tag{40}$$

where the dimensionless drag coefficient C is typically ~1. Thus, the characteristic distance required for drag to stop the cloud relative to the intercloud medium is $2\rho_c l/\rho$. Immediately after the viscous shock in the intercloud medium but before the cloud has had sufficient time to compress, this distance scale is comparable to the distance, $(n_s A)^{-1} = l/F_c$, required for cloud-cloud collisions to effect significant momentum transfer. However, the time required to compress the clouds to their downstream states is comparable to, if not shorter than, the most optimistic estimate of the deceleration time, and it is difficult to obtain from such crude arguments a precise numerical estimate of the deceleration distance of the cloud phase. Under certain circumstances the deceleration distance of the cloud phase may be so long (much greater, say, than 200 pc) that it is inappropriate to think in terms of a shock in this phase. In general, there may be a nonnegligible region (~200 pc in thickness) where the clouds move with a mean velocity which is substantially different from that of the intercloud phase.

The small-scale dynamics of the intercloud gas, in the vicinity of a cloud which plows through it, would also be rather complicated. The relative motion is initially supersonic

and would tend to produce a detached "bow shock" in the intercloud gas. The boundary-layer flow would have a line of separation, and there is a turbulent wake toward the rear of the cloud. Because of the inward acceleration of the heavy cloud gas by the light intercloud gas and because of the large shear across the interface, the interface layer itself tends to be unstable to a combination of the Rayleigh-Taylor and Kelvin-Helmholtz instabilities (Chandrasekhar 1961). The latter effect, coupled with the internal motions of the cloud and the additional compression of the intercloud gas behind the bow shock, may result in an ability for the clouds to act as "nucleation centers" for the transformation of intercloud material into cloud material.

The above discussions are, of course, highly speculative without further calculations. Some of these we hope to perform and present in a future paper.

VI. GRAVITATIONAL INSTABILITY OF CLOUDS AND STAR FORMATION

It is obvious from the discussion of § V*d* that the compression of clouds in the galactic shock layer is a fully dynamic affair. Of course, after $\sim 10^7$ years, a cloud which does not collapse gravitationally will inevitably reach a state of rough hydrostatic equilibrium consistent with the ambient downstream pressure. Therefore, it is possible to derive a *sufficient* condition for the gravitational collapse of a cloud by asking whether there exist *any* stable states of hydrostatic equilibrium accessible to the cloud for a given external pressure. Alternatively, for a given mass, we may ask for the range of external pressure to which the cloud may be subjected without *necessarily* causing gravitational collapse.

Such conditions have been given, for an isothermal gaseous sphere bounded by external pressure, by Spitzer (1965), who used the analysis given by Ebert, by Bonnor, and by McCrea and the tables tabulated by Chandrasekhar and Wares. A self-gravitating cloud heated by an external flux of energetic particles or photons will, however, not be isothermal since (i) the denser regions near the center must be cooler in thermal equilibrium than the more rarefied regions near the surface if the flux is uniform throughout the cloud, and (ii) the flux may be attenuated toward the center so that the central regions may be even cooler than estimated under the assumption of uniform flux. A cloud with such properties is easier to compress than an isothermal sphere; therefore, we expect it to be also more prone to gravitational instability.

a) Formal Development

To obtain a quantitative estimate of the modifications in the stability criterion, we assume, for simplicity, a uniform flux of low-energy cosmic rays throughout the cloud. The pressure P_c of the gas in the cloud then satisfies the approximate polytropic relation (6): $P_c = K\rho_c^{(N+1)/N}$, where $N < -1.0$ and $K = K(\zeta)$ a constant for given ζ. For generality, we leave N and K unspecified for the time being. The only novel feature in the analysis which follows is that the polytropic index N is negative. We remark that the continuous sequence of polytropes with positive indices to polytropes with negative indices occurs not through zero, but through infinity—i.e., through the isothermal sphere.

We assume that the self-gravitating cloud is nonmagnetic and nonrotating, and is in a state of hydrostatic equilibrium:

$$\frac{dP_c}{dr}(r) = -G\frac{M_c(r)\rho_c(r)}{r^2}, \qquad \frac{dM_c}{dr}(r) = 4\pi r^2\rho_c(r). \tag{41}$$

Except for a sign, we may use the standard transformation (Chandrasekhar 1939) to write

$$\rho_c(r) = \rho_c(0)\theta^N(\xi) \quad \text{with} \quad r = \alpha\xi \quad \text{and} \quad \alpha = \left\{-\frac{(N+1)K}{4\pi G[\rho_c(0)]^{(N-1)/N}}\right\}^{1/2}. \tag{42}$$

Equations (41) may then be manipulated to read

$$\frac{1}{\xi^2}\frac{d}{d\xi}\left(\xi^2\frac{d\theta}{d\xi}\right) = \theta^N , \tag{43}$$

which is to be solved under the initial conditions $\theta(0) = 1$ and $\theta'(0) = 0$.

The power series solution of equation (43) valid near the origin is obtained as

$$\theta(\xi) = 1 + \tfrac{1}{6}\xi^2 + \tfrac{1}{120}N\xi^4 + \ldots . \tag{44}$$

The full integration of equation (43) to reach the cloud's outer boundary at $\xi = \xi_1$ can be carried out numerically for general $N < -1$. For given K and N, the radius $R_c = \alpha\xi_1$, the surface pressure $P_c(R_c)$, and the total mass $M_c(R_c)$ can be found as functions of $\rho_c(0)$ in the usual way. We wish, however, to consider a sequence of clouds characterized by different values of ξ_1 with fixed total mass $M_c(R_c) = M$, rather than a sequence with fixed central density $\rho_c(0)$. Some straightforward algebra using the first relation of equation (41) and the relations of equation (42) yields the central density of the former sequence:

$$\rho_c(0) = \left(\frac{M}{4\pi}L_0^{-3}\right)[\xi_1^2\theta'(\xi_1)]^{-2N/(3-N)} . \tag{45}$$

In the above L_0 is a constant for given K, N, and M, and is the natural length scale in the problem:

$$L_0 = \left(\frac{M}{4\pi}\right)^{(1-N)/(3-N)}\left[-\frac{(N+1)K}{4\pi G}\right]^{N/(3-N)} . \tag{46}$$

In terms of this length scale, the radius, the surface pressure, and the average density of the cloud are given by the relations

$$R_c = L_0\xi_1[\xi_1^2\theta'(\xi_1)]^{(N-1)/(3-N)} , \tag{47a}$$

$$P_c(R_c) = K\left(\frac{M}{4\pi}L_0^{-3}\right)^{(N+1)/N}\theta^{N+1}(\xi_1)[\xi_1^2\theta'(\xi_1)]^{-2(N+1)/(3-N)} , \tag{47b}$$

$$\rho_{av} = \frac{3M}{4\pi R_c^3} = \left(\frac{M}{4\pi}L_0^{-3}\right)3\xi_1^{-3}[\xi_1^2\theta'(\xi_1)]^{-3(N-1)/(3-N)} . \tag{47c}$$

A measure of the degree of central concentration is given by the ratio of the central density $\rho_c(0)$ to the average density ρ_{av}:

$$\frac{\rho_c(0)}{\rho_{av}} = \frac{\xi_1^3}{3[\xi_1^2\theta'(\xi_1)]} . \tag{48}$$

From the series solution (44), we easily verify that the sequence of clouds characterized by small ξ_1 are nearly homogeneous, $\rho_c(0)/\rho_{av} \simeq 1$, and satisfy the approximate relation $P_c(R_c) = K\rho_{av}^{(N+1)/N}$. Evidently self-gravity plays an unimportant part in binding such clouds. The same is not true for larger values of ξ_1, as can be seen explicitly from table 3 for the specific case $N = -3.0$. In table 3 we have tabulated $\theta(\xi_1)$, $\xi_1^2\theta'(\xi_1)$, the dimensionless radius $\xi_1[\xi_1^2\theta'(\xi_1)]^{(N-1)/(3-N)}$, the dimensionless pressure $\theta^{N+1}(\xi_1)$ $[\xi_1^2\theta'(\xi_1)]^{-2(N+1)/(3-N)}$, and the ratio $\rho_c(0)/\rho_{av}$ as functions of the parameter ξ_1.

As ξ_1 increases from zero, the surface pressure increases, and the cloud becomes more centrally condensed as self-gravity becomes increasingly more important for binding the cloud. A critical value $\xi_1 = \xi_* = 3.80$ exists where the surface pressure reaches a maximum value 0.643. For pressures larger than this value there are no equilibrium con-

TABLE 3

PROPERTIES OF BOUNDED POLYTROPIC GAS SPHERES WITH $N = -3.0$

ξ_1	$\theta(\xi_1)$	$\xi_1^2\theta'(\xi_1)$	Radius	Pressure	$\rho_c(0)/\rho_{av}$
0.00.....	1.00	0.000	∞	0.000	1.00
0.20.....	1.01	0.003	10.52	0.019	1.02
0.40.....	1.03	0.020	5.36	0.071	1.05
0.60.....	1.06	0.065	3.70	0.145	1.11
0.80.....	1.10	0.144	2.91	0.228	1.19
1.00.....	1.15	0.259	2.46	0.310	1.29
1.20.....	1.20	0.409	2.18	0.383	1.41
1.40.....	1.26	0.591	1.99	0.445	1.55
1.60.....	1.32	0.801	1.85	0.495	1.70
1.80.....	1.38	1.04	1.76	0.535	1.88
2.00.....	1.45	1.29	1.69	0.565	2.07
2.20.....	1.51	1.56	1.63	0.589	2.27
2.40.....	1.58	1.85	1.59	0.606	2.49
2.60.....	1.64	2.15	1.56	0.619	2.73
2.80.....	1.70	2.46	1.54	0.628	2.97
3.00.....	1.77	2.78	1.52	0.634	3.23
3.20.....	1.83	3.11	1.50	0.639	3.51
3.40.....	1.89	3.45	1.49	0.641	3.79
3.60.....	1.95	3.80	1.48	0.643	4.09
3.80.....	2.00	4.16	1.47	0.643	4.40
4.00.....	2.06	4.52	1.46	0.643	4.72
4.20.....	2.12	4.89	1.46	0.642	5.05
4.40.....	2.17	5.26	1.45	0.641	5.40
4.60.....	2.23	5.64	1.45	0.639	5.75
4.80.....	2.28	6.03	1.45	0.637	6.12
5.00.....	2.33	6.42	1.45	0.636	6.49

figurations possible for the cloud. A second branch of equilibrium configurations exists at lower pressure with $\xi_1 > \xi_*$, but these are configurations of *unstable* equilibrium analogous to those well known for bounded isothermal spheres (see Spitzer 1968).

In summary, for given M, hydrostatic equilibrium with a given (dimensional) surface pressure $P_c(R_c)$ does not exist if $P_c(R_c)$ exceeds a certain critical value. Conversely, for given surface pressure $P_c(R_c)$ equal to the pressure P of the ambient intercloud medium, there exists a critical mass $M = M_*$ beyond which clouds are unstable toward gravitational collapse. A little manipulation yields M_* as

$$M_* = 4\pi \left(\frac{P}{K}\right)^{(3-N)/2(N+1)} \left[-\frac{(N+1)K}{4\pi G}\right]^{3/2} \frac{[\xi_*^2\theta'(\xi_*)]}{\theta^{(3-N)/2}(\xi_*)} . \tag{49}$$

An alternative form for equation (49) which is a little more convenient for computational purposes is obtained by using the relation $P = P_c(R_c) = K[\rho_c(R_c)]^{(N+1)/N}$ to yield

$$M_* = \mathfrak{M}(\xi_*)[\rho_c(R_c)]^{-2}(P/G)^{3/2} , \tag{50a}$$

$$\mathfrak{M}(\xi_*) = (4\pi)^{-1/2}[-(N+1)]^{3/2} \frac{\xi_*^2\theta'(\xi_*)}{\theta^{(3-N)/2}(\xi_*)} . \tag{50b}$$

In table 4 we have tabulated ξ_*, $\theta(\xi_*)$, $\xi_*^2\theta'(\xi_*)$, $\rho_c(0)/\rho_{av}$, $\rho_c(0)/\rho_c(R_c)$, and $\mathfrak{M}(\xi_*)$ for various values of N. For $P = P_{min}$ we use $N = -1.3$ (see table 1), and equation (50a) yields $M_* = 3000\ M_\odot$. For $P = P_{max}$ we use $N = -4.0$, and equation (50a) yields $M_* = 120\ M_\odot$. The above masses should be scaled by a factor $\lambda^{-1/2}$ if ζ is scaled by a factor λ. On the other hand, if there are effects (e.g., depletion of metals onto

TABLE 4

BOUNDED POLYTROPIC GAS SPHERES ON THE MARGIN OF STABILITY

N	ξ_*	$\theta(\xi_*)$	$\xi_*^2\theta'(\xi_*)$	$\rho_c(0)/\rho_{av}$	$\rho_c(0)/\rho_c(R_c)$	$\mathfrak{M}(\xi_*)$
-1.00.....	...	...	...	1.00	1.00	0.000
-1.50.....	5.70	3.35	16.0	3.86	6.13	0.105
-2.00.....	4.78	2.63	8.91	4.08	6.90	0.225
-2.50.....	4.22	2.25	5.85	4.28	7.58	0.327
-3.00.....	3.80	2.00	4.16	4.40	8.06	0.411
-3.50.....	3.50	1.84	3.16	4.52	8.52	0.482
-4.00.....	3.26	1.73	2.50	4.62	8.86	0.541
$-\infty$.....	...	...	...	5.78	14.3	1.18

interstellar grains, internal circulation or turbulence, internal magnetic fields) which increase the "effective pressure" inside the cloud above the normal gas kinetic pressure, we may crudely account for these effects by increasing the value K in equation (49) without increasing P. The masses M_* which are marginally stable then increase as $K^{2N/(N+1)}$. The exponent $2N/(N+1) > 2$ for $N < -1$, so the masses M_* are very sensitive to such an an increase of the "effective internal pressure." One major difference between Roberts's (1969) estimate of the critical mass for collapse and ours is that we do not include the effects of internal turbulence in the support of the cloud. Rather, we would follow Stone (1970) in attributing the observed internal velocity dispersion of clouds mostly to the bulk expansional motion that eventually follows the collision of two clouds.

b) Comments

Our values for the critical masses for stability are, of course, approximate in the sense that the polytropic law (6) applies with only approximate validity for the cloud phase. Nevertheless, a comparison of tables 1 and 4 shows the true polytropic index not to vary appreciably over the range of densities from surface to center encountered in a cloud on the margin of stability. Furthermore, exact calculations of the critical masses (with $P = P_{\min}$ and $P = P_v$) by Penston and Brown for an equilibrium P-ρ relation which is similar to ours yield results in rough agreement with ours. We have not felt exact calculations to be worthwhile because the thermal-balance calculations are themselves no longer reliable in the central regions of a marginally stable cloud when $P = P_{\max}$. There, the atomic density would reach $\sim$400 cm^{-3} and cooling by radiation in the fine-structure lines ("forbidden" transitions) would be quenched by collisional deexcitation.

In any case, interpreted literally our results imply that no clouds in excess of 3000 $M_\odot$ can exist anywhere along a circle of radius 10 kpc in the Galaxy, whereas a galactic shock wave will necessarily trigger gravitational collapse in all preexisting clouds with masses greater than 120 $M_\odot$. Collapse may be induced in even smaller masses because transient pressures higher than $P_{\max}$ can be reached inside the galactic shock layer. Of course, these estimates would be considerably modified if ζ has a different value in spiral arms than between spiral arms or if a cloud has means of support other than its internal gas kinetic pressure.

Our purpose here is not to give definitive limits for the masses of clouds which are marginally stable, but to emphasize that a "mere" increase of the external pressure by a factor 5 suffices to reduce the critical cloud mass for stability by a factor $25 = 5^2$. This latter factor is to be contrasted with the factor $5^{1/2} = 2.2$ which would have resulted if the clouds were compressed isothermally (cf. eq. [49] in the limit $N \to -\infty$). The very substantial decrease in the threshold for gravitational collapse is explained by the

greater compressibility of the gas and the nonlinear nature of self-gravity. This facet of our model, coupled with the concept of galactic shocks, goes a long way toward explaining why the regions of substantial star formation are delineated very sharply in certain external galaxies.

c) *The Problem of Fragmentation*

Finally, we remark that the gravitational collapse of clouds of a few thousand solar masses would not, of course, lead to the formation of a single star. (The clouds of ~ 100 $M_\odot$ offer fewer problems in this respect.) We expect the collapse of such clouds or of complexes of such clouds to lead to the formation of a star cluster. The fragmentation by gravitational instability of such a gas cloud is an old and controversial issue (cf. the reviews by Layzer 1964 and by Hunter 1967). An essential difficulty is that the timescale for the formation of gravitational subcondensations is of the same order, $\sim 10^7$ years, as that for the free-fall collapse of the cloud as a whole.[8]

We wish, here, merely to add two possibilities to the list of those which have been discussed previously. Fragmentation may proceed if large fluctuations in density (subclouds) are initially present in the interior of the cloud because a (hypothetical) third phase involving molecules and/or dust can exist in rough pressure equilibrium with the usual cloud phase. It may also proceed if regions slightly denser than average achieve a separate identity by thermal processes on a timescale of $\sim 10^4$ years because such regions will cool more rapidly than the average after the entire cloud has been subjected to an internal shock (see § V*d*). The latter mechanism is closely related to the thermal instabilities discussed by Field (1965), Hunter (1966), and McCray and Schwarz (1971). In any case, the self-gravity of the subcloud would eventually be required to separate it permanently from the rest of the collapsing cloud.

VII. DISCUSSION

We have seen that the large-scale dynamics of a two-component model of the interstellar medium is similar in many respects to that of the one-component model. However, by introducing the two-component model, we can present a clear physical basis for the mechanisms of the production of the shock and of the compression of the clouds. We also acquire the possibility of transitions between the two phases.

In particular, phase transitions of cloud material into intercloud material results in maintaining the pressure of the regions between spiral arms at the lowest possible level consistent with the existence of clouds. Thus, the detection of interstellar clouds in these regions by 21-cm absorption should be very difficult. The last comment is related to a more general one. While it is possible to interpret 21-cm emission profiles without examining the thermodynamics of the medium in detail (provided one accounts for the effects of "velocity crowding" and "velocity stretching"), it is dangerous to do the same for the absorption profiles. The reason is that the temperature of the clouds, in our model, varies systematically from $T \sim 30°$–$50°$ K in spiral arms to $T \sim 200°$ K between spiral arms.

An important question not answered in this paper but under investigation by P. Woodward (private communication) is the relation between flow solutions that contain shocks and those that do not. Shock waves are essential only for defining a *narrow lane* of star formation—not for the *existence* of the process itself. The thick spiral arms and somewhat mottled appearance of some Sc galaxies may well be associated with an underlying compression of the interstellar medium which is more smooth than that obtained in the usual shock picture.

[8] Numerical calculations by Larson and by others (see the review by Penston 1971) have partially alleviated this problem by showing that a centrally condensed core, at least, will collapse well before the envelope has contracted appreciably. Such calculations, however, start with a fairly dense object ($n \sim 10^5$ cm^{-3}) of stellar mass as a postulate and do not indicate how a large interstellar cloud will fragment into many smaller subunits which can be reasonably identified with protostars.

Apart from the above question and the calculation of the (presumably small) differences of velocities between the two phases, our work on the large-scale problem is fairly complete. The same is not true for the small-scale phenomena. We have given in this paper a criterion for the gravitational collapse of an interstellar cloud. In particular, for that rare nonrotating 120 $M_\odot$ cloud devoid of internal turbulence and magnetic fields, the shock mechanism may be capable of producing the formation of a single star from its centrally condensed core. Left unsolved is the problem of the dynamical collapse of a much larger cloud under sudden loading and the (possibly thermal) mechanisms by which it fragments into a cluster of protostars. We have also not investigated the internal structure of a shock later which includes the interactions between the two phases. Such an investigation may show that the relaxation layer associated with the cloud phase is, under certain circumstances, too thick to justify treating it as a discontinuity in the large-scale flow.

The present work examines the effects of dynamical forcing of the interstellar medium by a background spiral gravitational field. What about the reaction of the medium back onto the spiral structure? Here, the key role played by galactic shocks is probably one of dissipation. The nonclosure of the streamlines discussed in § IV*c* indicates that the periodically shocked gas tends to lose energy and angular momentum. If this energy and angular momentum is gained by the spiral wave, instead of being radiated away, say, or going toward rearranging the basic state, then the process may damp an otherwise steady wave, or help to limit the amplitude of a slowly growing wave. This possibility arises because the spiral waves of the type considered by Lin and Shu (1964, 1966, 1972) contain *negative* densities of energy and angular momentum (cf. Toomre 1969; Shu 1970; Kalnajs 1970). However, the physical effect is probably small, and a more careful analysis of this point is required to determine whether the shock mechanism does play an important role in the problem of the permanence of spiral structure.

Finally we remark that our calculations have been formally performed for a steady state. Input concerning the mass distribution of clouds and the birth function of stars would allow the use of these calculations to estimate the chemical evolution of the Galaxy.

We wish to acknowledge fruitful discussions with Professors C. C. Lin, L. Woltjer, George Field, Stephen Strom, Alar Toomre, James Mark, Deane Peterson, and Richard McCray. The numerical calculations were carried out at the Computation Centers at Stony Brook, at M.I.T., at the NASA Institute for Space Studies, and at Charlottesville. The work of F. H. S. was supported in part by NSF grants GP-13061 (Stony Brook) and GP-22720 (M.I.T.); that of C. Y., by the Faculty Research Award No. 1349 (CUNY); that of D. W. G., by NSF grant GP-18476; and that of W. W. R., by NSF grant GP-25618.

APPENDIX A

KINETIC DESCRIPTION OF THE CLOUDS

Let $f(\boldsymbol{x}, \boldsymbol{v}, M, t)d^3xd^3vdM$ be the number of clouds with mass between M and $M + dM$, with position in the volume element d^3x centered about $\boldsymbol{x}$, and with velocity in the volume d^3v of velocity space centered about $\boldsymbol{v}$. The dynamic evolution of f is governed by the kinetic equation

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial \boldsymbol{x}}\cdot(\boldsymbol{v}f) + \frac{\partial}{\partial \boldsymbol{v}}\cdot\left[\left(-\frac{\partial \mathcal{U}}{\partial \boldsymbol{x}} + \frac{\boldsymbol{F}_{\rm ic}}{M} + \frac{\dot{M}}{M}\boldsymbol{v}_{\rm rel}\right)f\right] + \frac{\partial}{\partial M}(\dot{M}f)$$
$$= \left(\frac{\delta f}{\delta t}\right)_{\rm coll} + \left(\frac{\delta f}{\delta t}\right)_{\rm acc}. \tag{A1}$$

In the above $\mathcal{U}$ is the gravitational potential of the Galaxy, $\boldsymbol{F}_{ic}$ is the force exerted on a cloud by the intercloud medium, $\dot{M}$ is the rate of change of the mass of a cloud due to phase transition, and $\boldsymbol{v}_{rel}$ is the velocity of the material involved in the phase transition relative to the velocity of the cloud. The two terms on the right-hand side represent, respectively, the effects of (inelastic) cloud-cloud collisions and of (random) mechanisms of cloud accelerations by small-scale processes (supernovae explosions, expanding H II regions, etc.). We have ignored the net sink of clouds due to star formation.

To make contact with the usual model of "standard clouds" we introduce the following "moments" of the distribution function:

$$\begin{pmatrix} n_s \\ n_s\langle M\rangle \\ n_s\langle M\rangle \boldsymbol{u}_c \end{pmatrix} = \int \begin{pmatrix} 1 \\ M \\ M\boldsymbol{v} \end{pmatrix} f d^3v dM \,, \tag{A2}$$

where the various symbols are defined in § III. In this paper we have preferred to use the combination $F_c\langle\rho_c\rangle$ in lieu of $n_s\langle M\rangle$ (see eq. [8]).

We assume that cloud-cloud collisions and the random mechanisms of cloud accelerations do not lead to any systematic changes locally of either the total mass contained in the form of clouds or the total vector momentum possessed by the clouds. Thus, the right-hand side of equation (A1) has the properties

$$\int \begin{pmatrix} M \\ M\boldsymbol{v} \end{pmatrix} \left[\left(\frac{\delta f}{\delta t}\right)_{coll} + \left(\frac{\delta f}{\delta t}\right)_{acc}\right] d^3v dM = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{A3}$$

These important properties allow us to write the first two "moment equations" in the usual way as the conservation relations (10a) and (10b) for mass and momentum. Because the collisions are inelastic, the analog of the usual energy equation cannot be written down without a detailed knowledge of the structure of the right-hand terms in equation (A1). Furthermore, if clouds coalesce or splatter upon collision, the conservation of mass is no longer synonymous with the conservation of number.

In equations (10a) and (10b) we have written $n_s\langle\dot{M}\rangle$ and $n_s\langle\dot{M}\boldsymbol{v}'\rangle$ for the integrals

$$\begin{pmatrix} n_s\langle\dot{M}\rangle \\ n_s\langle\dot{M}\boldsymbol{v}'\rangle \end{pmatrix} = \int \begin{pmatrix} \dot{M} \\ \dot{M}(\boldsymbol{v} + \boldsymbol{v}_{rel}) \end{pmatrix} f d^3v dM \,, \tag{A4}$$

whereas in equation (12) we have implicitly assumed that $\boldsymbol{v}_{rel}$ is zero if cloud material is being transformed into intercloud material and is $(\boldsymbol{u} - \boldsymbol{v})$ if intercloud material is being transformed to cloud material by accretion onto the preexisting clouds. The preconceived notion that clouds act as "nucleation centers" for the latter transition is implicit in our notation in § III; however, it does not constitute an essential ingredient in our macroscopic treatment of phase transitions.

In equation (10b) we have also assumed f to be isotropic in $(\boldsymbol{v} - \boldsymbol{u}_c)$ so that the stress tensor is diagonal:

$$\int (\boldsymbol{v} - \boldsymbol{u}_c)(\boldsymbol{v} - \boldsymbol{u}_c) M f d^3v dM = F_c\langle\rho_c\rangle c^2 \boldsymbol{I} \,, \tag{A5}$$

where $\boldsymbol{I}$ is the unit dyadic. In theory the determination of the local value of c should come from the balancing of the two terms on the right-hand side of equation (A1) because these two terms have associated timescales which are at least one order of magnitude smaller than those of the terms on the left-hand side. For practical calculations we have taken c to be a constant, 8 km s^{-1}. We further remark that the assumption of isotropy of f in $(\boldsymbol{v} - \boldsymbol{u}_c)$ must break down in the vicinity of a galactic shock layer where $\boldsymbol{u}_c$ changes appreciably over one mean free path for cloud-cloud collisions. The inclusion of this effect in a fluid description would presumably lead to terms analogous to those of

"viscous stress"; however, the proper treatment of the structure of the "viscous shock" in the cloud phase may require a kinetic description.

In equation (10b) we have written the average body force exerted on the clouds by the intercloud medium as

$$\int \boldsymbol{F}_{ic} f d^3v dM = -F_c \nabla P - n_s[C\pi\langle R_c\rangle^2(1 - F_c)\tfrac{1}{2}\rho|\boldsymbol{u}_c - \boldsymbol{u}|(\boldsymbol{u}_c - \boldsymbol{u})], \quad \text{(A6)}$$

where the term in brackets represents the drag on an average cloud because of the relative motion $(\boldsymbol{u}_c - \boldsymbol{u})$ in an intercloud medium of density ρ which occupies a fraction $(1 - F_c)$ of the volume of interstellar space. If we compare the expressions giving the drag force in equations (10b) and (A6), we find that the coefficient D is indeed given by equation (11).

APPENDIX B

SIMPLE MODEL OF THE VERTICAL STRUCTURE

The noncircular flow induced by the spiral gravitational field of our Galaxy will not produce fluid velocities u_z in the vertical direction which are in excess of 2–3 km s^{-1}. The rms random velocity of the clouds and the thermal speed of the intercloud medium are $\gtrsim$8 km s^{-1}. Since pressure and inertial effects are proportional to the square of the velocities, the vertical structures of the two phases, as determined by the z-component of the momentum equations (10b) and (13b), nearly satisfy the conditions of local hydrostatic balance:

$$\frac{\partial}{\partial z}(F_c\langle\rho_c\rangle c^2) = -F_c\langle\rho_c\rangle\frac{\partial \mathcal{U}}{\partial z}, \qquad \frac{\partial P}{\partial z} = -\rho\frac{\partial \mathcal{U}}{\partial z}. \quad \text{(B1)}$$

To integrate these equations analytically, we make the following ad hoc assumptions. We assume that c is independent of z. We denote any variable evaluated at $z = 0$ by a superscript zero and further suppose that ρ and ζ have identical distributions ψ in z:

$$\rho = \rho^0\psi, \qquad \zeta = \zeta^0\psi, \quad \text{(B2)}$$

where $\psi = 1$ at $z = 0$. (Of course, the assumption that the primary flux of ionization falls off with increasing z would be incorrect even qualitatively if the sources of ionization were extragalactic.) Because of the scaling law (5) for P_{eq}, equations (14) and (7) imply

$$P = P^0\psi, \qquad \langle\rho_c\rangle = \langle\rho_c\rangle^0\psi, \quad \text{(B3)}$$

i.e., the mean temperatures of the two phases are independent of z. The integrations of equations (B1) now yield

$$F_c\langle\rho_c\rangle = F_c{}^0\langle\rho_c\rangle^0 \exp\left(-\frac{\mathcal{U} - \mathcal{U}^0}{c^2}\right), \qquad \psi = \exp\left(-\frac{\mathcal{U} - \mathcal{U}^0}{P^0/\rho^0}\right). \quad \text{(B4)}$$

The source for the gravitational potential is primarily the disk stars. Since c^2 and P^0/ρ^0 are small in comparison with the mean square dispersive velocities of the disk stars and since $\partial\mathcal{U}/\partial z = 0$ at $z = 0$ because the latter is assumed to be a plane of reflection symmetry, $z = 0$ is a saddlepoint for the distributions (B4). In the lowest order of approximation, we may replace $\mathcal{U} - \mathcal{U}^0$ by the first nonvanishing term in its Taylor series expansion about $z = 0$:

$$\mathcal{U} - \mathcal{U}^0 = \left(\frac{\partial^2\mathcal{U}}{\partial z}\right)_{z=0}\frac{z^2}{2}. \quad \text{(B5)}$$

Thus, the distributions (B4) are approximately Gaussians in z with effective thickness h_c and h defined through equations (16). The integrations of equations (10) and (13)

over z may now be carried out asymptotically by the method of steepest descent by treating $\mathcal{U}$, $\boldsymbol{u}_c$, $\boldsymbol{u}$, $\langle M\rangle$, and $\langle \dot{M}\rangle$ as functions which are slowly varying over distances of h_c or h. The integrated forms of equations (10) and (13), written to lowest order, are identical to their original forms except that everywhere we see $F_c\langle\rho_c\rangle$, n_s, $(1 - F_c)\rho$, or $(1 - F_c)\nabla P$; we replace it by $F_c{}^0\langle\rho_c\rangle^0 h_c$, $n_s{}^0 h_c$, $(1 - F_c{}^0)\rho^0 h$, or $(1 - F_c{}^0)\nabla(P^0 h)$; whereas everywhere we see $\mathcal{U}$, $\boldsymbol{u}_c$, $\boldsymbol{u}$, $\langle M\rangle$, or $\langle \dot{M}\rangle$, we replace it by $\mathcal{U}^0$, $\boldsymbol{u}_c{}^0$, $\boldsymbol{u}^0$, $\langle M\rangle^0$, or $\langle \dot{M}\rangle^0$.

Two terms in equation (10b), $F_c\nabla P$ and $D(\boldsymbol{u}_c - \boldsymbol{u})$, give potential difficulty because equations (B3) and (B4) imply that F_c is a slowly varying function of z only if c^2 should happen to be nearly equal to P^0/ρ^0. Nevertheless, we shall replace $F_c\nabla P$ by $F_c{}^0\nabla(P^0 h)$ on the ground that the former is, in any case, small in comparison with $\nabla(F_c\langle\rho_c\rangle c^2)$; and we shall replace D by $D^0 h$ on the ground that any error incurred in this way can be absorbed in a redefinition of C. If we now drop the cumbersome superscript zero notation and write the resulting equations in the frame which rotates with angular velocity Ω_p, we obtain equations (17).

APPENDIX C

THE CALCULATION OF THE LARGE-SCALE GALACTIC FLOW

The most fundamental aspect of shocked flows is the existence of a region of supersonic flow and a region of subsonic flow. The passage from supersonic flow to a region of subsonic flow is achieved via a shock. If the streamlines are (nearly) closed, there must also be a passage from subsonic flow to supersonic flow through a sonic point.

A smooth passage through a sonic point can occur only under special conditions—e.g., at a "throat" in nozzle flow. As is well known from Parker's solution for a steady solar wind (see Parker 1963), the role of a "throat" often appears in a modified form when external forces are present. From equation (30b), we see that $\partial u_\eta/\partial\eta$ remains finite when u_η equals a only if $(2\Omega u_{\xi 1} + g_1)$ is zero. Indeed, for given location (ϖ, $\eta = \eta_{SP}$) of the sonic point and for given ρ (i.e., for given a) at the sonic point, a series can be developed in powers of $(\eta - \eta_{SP})$ to obtain the solution of the system (30) which is regular in the neighborhood of the sonic point. The values chosen for η_{SP} and ρ at the sonic point can be adjusted later to satisfy the conditions that the flow be doubly periodic and that the net phase transformation in a complete circuit (actually, in a 180° circuit) around the Galaxy be zero. The extra degree of freedom, the choice of f_c (or F_c) at the sonic point, is at our disposal to fit the observations pertaining to the amount of neutral hydrogen present in the Galaxy.

The values of ρ and f_c at the sonic point carry no special physical significance. We find it conceptually useful to relate these values to the "average" values for a streamtube of average radius ϖ. Using the word "average" loosely, we denote ρ_0 as the average density of the intercloud gas, F_{c0} as the average fraction of volume occupied by the clouds, and relate the conditions at the sonic point through the mass fluxes

$$(\rho a)_{SP} = \rho_0 u_{\eta 0}\,, \qquad (f_c\langle\rho_c\rangle a)_{SP} = f_{c0}\langle\rho_c\rangle_0 u_{\eta 0}\,, \tag{C1}$$

where $u_{\eta 0}$ is given through equation (32a). By using the mass flux to define the various "average" quantities, we have automatically weighted the "averaging" process by the volume; i.e., the averaging is more properly regarded as one over the *streamtube* than as one along the *streamline*. From ρ_0, we further define the "average pressure" P_0 through the equilibrium P-ρ relation.

Consider now a particular guess for η_{SP} and P_0. As outlined above, the sonic point is a natural point to start the numerical integration of the flow—forward in η to obtain the supersonic branch, and backward to obtain the subsonic branch. These two branches are later to be joined by a shock; we now describe the procedure for determining the location of the shock as well as for determining the other characteristics of the flow.

As we integrate *backward* along the subsonic branch, the flow is compressional and the pressure increases steadily. Note, however, that P can never increase to $P_{\max}$ in subsonic flow because $a^2 = (\partial P_{eq}/\partial \rho)_\zeta$ is zero at $P = P_{\max}$, and the gradient of u_η as given by equation (30b) would become infinite before $P_{\max}$ is actually reached. Thus, if P reaches or exceeds $P_{\max}$ anywhere in the flow, it can do so only within the shock layer. (This statement is not true if the flow is entirely supersonic along the whole streamline and there are no shocks.) If phase transition is present in the shock layer, the postshock pressure $P(2)$ must be near but slightly less than $P_{\max}$ (see eq. [21]). The amount that $P(2)$ is less than $P_{\max}$ is, of course, not precisely determined simply because there is never a clear demarcation between the shock layer and the rest of the flow. We must be satisfied that our numerical results are not sensitive to the exact choice adopted for $P(2)$ when phase transitions are present. When phase transitions are absent, the postshock pressure may have any value less than $P_{\max}$.

To find an acceptable solution in actual practice, we use the following convenient device. We use the exact P-$\langle \rho_c \rangle$ relation for the clouds but perform the calculations for the intercloud medium using the polytropic P-ρ relation (6) *even beyond* $P = P_{\max}$ if need be. Having obtained the subsonic branch, we perform the supersonic integration by marching *forward* from the sonic point using equations (30). We switch to equations (33) if P tries to decrease below $P_{\min}$ and back to equations (30) if the flow becomes compressional (see § IV*b*). At each step of the supersonic integration we artificially compute the postshock conditions that would result from the jump conditions (26) *were the gas to shock at that point*. (The value used for δ_m is obtained by subtracting $(\rho a)_{SP}$ from the local value of ρu_η.)

The artificial $u_\eta(2) = u_\perp(2)$ computed in this way is plotted for each local value of u_ξ in the (u_η, u_ξ)-plane together with the results of the subsonic and supersonic integrations (see fig. 5). The intersection of the curve obtained for the jump conditions and that for the subsonic branch singles out the conditions under which the gas can "shock" in this scheme. However, in general, the difference in azimuthal angle reached by the super-

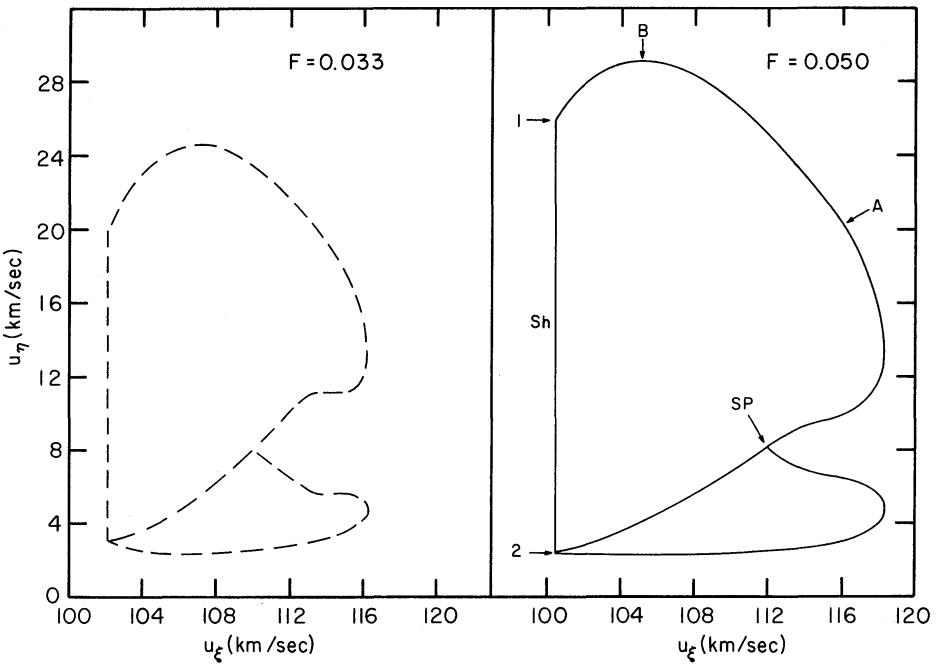


FIG. 5.—The procedure used to locate the galactic shocks. See text in Appendix C for explanations. The labeling used here is consistent with that used in fig. 2.

sonic flow and that reached by the subsonic flow will not be 180° (for a doubly periodic solution), nor will $P(2)$ be near $P_{\max}$ if phase transitions are present. Such a trial integration would not constitute an acceptable solution—especially if $P(2)$ is greater than $P_{\max}$. We successively choose different starting values of η_{SP} and P_0 until both conditions, double periodicity and $P(2) = P_{\max}$, are satisfied. An efficient iteration procedure, based on a two-dimensional analog of Newton's method for finding roots, can find in four or five tries a solution which satisfies both conditions to an accuracy of 1 percent or better.

APPENDIX D

THE LIKELIHOOD OF THE EXISTENCE OF A SONIC POINT

As described in the previous Appendix, the possibility of finding streamline solutions which contain shocks depends intimately on the existence of a sonic point. For small relative strengths F of the spiral gravitational field, the velocity component u_η cannot be expected to vary by more than $\pm u_{\eta 0}$ about $u_{\eta 0}$. Thus u_η has a chance of reaching a only if the "unperturbed" velocity $u_{\eta 0}$ is, to begin with, not too far from typical values for a.

Since $u_{\eta 0}$ is not related to a in any obvious manner, the condition $u_{\eta 0} \sim a$ (in practice, $u_{\eta 0} \gtrsim a$) seems at first sight unlikely to be satisfied in general. In fact, direct examination of figure 6 in Shu *et al.* (1971) shows $u_{\eta 0}$ (denoted $w_{\perp 0}$ by Shu *et al.*) to be 10–20 km s^{-1} over a wide range of radii for three galaxies—M33, M51, and M81—which differ substantially in kinematic properties.

The explanation rests with the (asymptotic) dispersion relation for (linear) density waves and with the recognition that $u_{\eta 0}$ represents the phase velocity of the wave with respect to the circular motion of the matter. The "short spiral waves," when sustained primarily by the disk stars, are exactly those for which this phase velocity is comparable to (but usually a factor 2–3 smaller than) the stellar rms velocity dispersion $\langle c_\varpi^2 \rangle^{1/2}$ in the radial direction. For a wide class of mass models for normal spirals, gravitational stability of the galactic disk requires $\langle c_\varpi^2 \rangle^{1/2} \gtrsim 30$ km s^{-1}, whereas, independent of ζ, thermal stability of the intercloud medium requires $a \gtrsim 7$ km s^{-1}. This set of circumstances virtually guarantees $u_{\eta 0}$ in normal spiral galaxies to have values which makes transonic flow a likely condition. (Short spiral waves of small amplitude supported *entirely* by the self-gravity of the *gas* lead to values of $u_{\eta 0}$ which are *subsonic* over the entire disk. This may represent an important difference between Vandervoort's 1971 model and ours.)

REFERENCES

Bottcher, C., McCray, R. A., Jura, M., and Dalgarno, A. 1970, *Ap. Letters*, **6**, 237.
Chandrasekhar, S. 1939, *An Introduction to the Study of Stellar Structure* (New York: Dover Publishing Co.).
———. 1961, *Hydrodynamic and Hydromagnetic Stability* (Oxford: Clarendon Press).
Chernyi, G. G. 1961, *Introduction to Hypersonic Flow* (New York: Academic Press).
Drew, D. 1971, *Studies Appl. Math.*, **50**, 133.
Eddington, A. S. 1926, *Proc. Roy. Soc.* (*London*), *A*, **111**, 424.
Field, G. B. 1965, *Ap. J.*, **142**, 531.
Field, G. B., Goldsmith, D. W., and Habing, H. J. 1969, *Ap. J.* (*Letters*), **155**, L149.
Field, G. B., Rather, J. D. G., Aanestad, P. A., and Orszag, S. A. 1968, *Ap. J.*, **151**, 953.
Fujimoto, M. 1966, in *I.A.U. Symp. No. 29*, p. 453.
Goldsmith, D. W. 1970, *Ap. J.*, **161**, 41.
Goldsmith, D. W., Habing, H. J., and Field, G. B. 1969, *Ap. J.*, **158**, 173.
Habing, H. J., and Goldsmith, D. W. 1971, *Ap. J.*, **166**, 525.
Hjellming, R., Gordon, C. P., and Gordon, K. J. 1969, *Astr. and Ap.*, **2**, 202.
Hunter, C. 1967, in *Relativity Theory and Astrophysics 2. Galactic Structure*, ed. J. Ehlers (Providence: American Mathematical Society), p. 169.
Hunter, J. H. 1966, *M.N.R.A.S.*, **133**, 239.
Jager, C. de. 1970, *Nature*, **225**, 622.
Kalnajs, A. J. 1970, in *The Spiral Structure of Our Galaxy*, ed. W. Becker and G. Contopoulos (Dordrecht: D. Reidel Publishing Co.), p. 314.

Layzer, D. 1964, *Ann. Rev. Astr. and Ap.*, **2**, 341.
Lin, C. C., and Shu, F. H. 1964, *Ap. J.*, **140**, 646.
———. 1966, *Proc. Nat. Acad. Sci.*, **55**, 229.
———. 1972, *Theory of Galactic Spirals* (Cambridge: M.I.T. Press), in preparation.
Lin, C. C., Yuan, C., and Shu, F. H. 1969, *Ap. J.*, **155**, 721.
McCray, R. A., and Schwarz, J. 1971, in *The Gum Nebula and Related Problems* (New York: NASA).
Parker, E. N. 1963, *Interplanetary Dynamical Processes* (New York: Interscience Publishers).
Penston, M. V. 1970, *Ap. J.*, **162**, 771.
———. 1971, *Contemp. Phys.*, **12**, 379.
Penston, M. V., and Brown, F. E. 1970, *M.N.R.A.S.*, **150**, 373.
Peterson, D. M., and Strom, S. E. 1969, *Ap. J.*, **157**, 1341.
Quirk, W. J. 1971, preprint.
Roberts, W. W. 1969, *Ap. J.*, **158**, 123.
Roberts, W. W., and Yuan, C. 1970, *Ap. J.*, **161**, 887.
Schmidt, M. 1965, in *Galactic Structure*, ed. A. Blaauw and M. Schmidt (Chicago: University of Chicago Press), p. 513.
Shu, F. H. 1970, *Ap. J.*, **160**, 99.
Shu, F. H., Stachnik, R. V., and Yost, J. C. 1971, *Ap. J.*, **166**, 465.
Silk, J., and Werner, M. W. 1969, *Ap. J.*, **158**, 185.
Simonson, S. C. 1970, *Astr. and Ap.*, **9**, 163.
Spitzer, L. 1968, in *Nebulae and Interstellar Matter*, ed. L. Aller and B. Middlehurst (Chicago: University of Chicago Press), p. 1.
Spitzer, L., and Scott, E. H. 1969, *Ap. J.*, **158**, 161.
Stone, M. E. 1970, *Ap. J.*, **159**, 293.
Toomre, A. 1969, *Ap. J.*, **158**, 899.
Vandervoort, P. O. 1971, *Ap. J.*, **166**, 37.
Whitney, C. A., and Skalafuris, A. J. 1963, *Ap. J.*, **138**, 200.
Yuan, C. 1969, *Ap. J.*, **158**, 889.
Zel'dovich, Ya. B., and Pikel'ner, S. B. 1969, *Zh. Eksper. i Teoret. Fiz.*, **56**, 310.
Zel'dovich, Ya. B., and Raizer, Yu. B. 1966, *Physics of Shock Waves and High Temperature Hydrodynamic Phenomena* (New York: Academic Press).